

The Plaid Model

by

Richard Evan Schwartz

Contents

0.1	Part 1: The Plaid Model	14
0.2	Part 2: The Plaid Master Picture Theorem	15
0.3	Part 3: The Arithmetic Graph	16
0.4	Part 4: The Quasi Isomorphism Theorem	16
0.5	Part 5: The Distribution of Orbits	18
0.6	Companion Program	20
I	The Plaid Model	21
1	Two Definitions of the Plaid Model	22
1.1	Basic Parameters	22
1.2	Six Families of Lines	22
1.3	Capacity, Mass, and Sign	24
1.4	First Definition of the Plaid Model	25
1.5	Second Definition of the Plaid Model	28
1.6	The Directed Version	30
2	Basic Properties of the Model	32
2.1	A Characterization of the Masses and Capacities	32
2.2	Symmetries	33
2.3	Symmetry and Direction	35
2.4	The Number of Intersection Points	36
2.5	Capacity and Mass	37
3	Using the Model	39
3.1	The Big Polygon	39
3.2	Hierarchical Information	41
3.3	Subdivision Algorithm	45
3.4	Grid Lines as Barriers	47
4	Three Dimensional Construction	50
4.1	Remote Adjacency	50
4.2	Horizontal Particles	51
4.3	Vertical Particles	53
4.4	Stacking the Blocks	55
4.5	Pixellated Spacetime Diagrams	57

4.6	Spacetime Plaid Surfaces	59
4.7	The Simplest Example	61
4.8	A Remark about Rescaling Limits	62
5	Directed Spacetime Diagrams	63
5.1	The Basic Definition	63
5.2	The Curve Turning Process	64
5.3	Two Exclusion Principles	65
5.4	Proof of the Curve Turning Theorem	66
6	Connection to the Truchet Tile System	70
6.1	Truchet Tilings	70
6.2	The Truchet Comparison Theorem	72
6.3	A Result from Elementary Number Theory	75
6.4	Proof of the Truchet Comparison Theorem	77
6.5	Symmetry of the Horizontal Diagrams	80
6.6	Symmetry of the Vertical Diagrams	81
II	The Plaid Master Picture Theorem	83
7	The Plaid Master Picture Theorem	84
7.1	The Spaces	84
7.2	The Partition	85
7.3	The Map	89
7.4	Three Dimensional Interpretation	91
7.5	The Main Result	92
7.6	The Undirected Result	93
8	The Images of Particles	94
8.1	Overview	94
8.2	The Vertical Case	95
8.3	The Horizontal P Case	95
8.4	The Horizontal Q Case	96
9	Proof of the Main Result	98
9.1	Symmetric Instances	98
9.2	Sharp Containers	99
9.3	Prism Structure	101

9.4	The Vertical Case	103
9.5	The Horizontal Case	105
10	Proof of the Vertical Lemma	107
10.1	Outline	107
10.2	Using Symmetry	107
10.3	Translating the Picture	108
10.4	Forgetting the Directions	109
10.5	Some Useful Formulas	110
10.6	The Undirected Result	111
10.7	Determining the Directions	113
11	Proof of The Horizontal Lemma	114
11.1	Using Symmetry	114
11.2	Translating the Picture	115
11.3	A Technical Lemma	116
11.4	The Undirected Result	116
11.5	Determining the Directions	117
III	The Graph Master Picture Theorem	119
12	The Arithmetic Graph	120
12.1	Special Orbits and the First Return Map	120
12.2	The Arithmetic Graph	121
12.3	The Canonical Affine Transformation	123
12.4	Geometry of the Graph Grid	124
13	Graph Master Picture Theorem	129
13.1	Statement of the Result	129
13.2	Pulling Back the Maps	131
13.3	Further Discussion	131
13.4	The Fundamental Polytopes	133
14	Pinwheels and Quarter Turn Systems	135
14.1	Overview	135
14.2	The Pinwheel Map	136
14.3	Outer Billiards and the Pinwheel Map	138
14.4	Quarter Turn Compositions	139

14.5	The Pinwheel Map as a QTS	140
14.6	The Case of Kites	143
15	The General Compactification	145
15.1	Affine Pets Redefined	145
15.2	The Map	146
15.3	Extending The Component Maps	147
15.4	The Composition	149
15.5	Double Lattice PETs	150
15.6	The Structure Theorem	151
15.7	The General Arithmetic Graph	152
15.8	The Case of Kites	153
16	Proof of the Structure Theorem	155
16.1	The Singular Directions	155
16.2	The First Parallelotope	157
16.3	The Second Parallelotope	159
16.4	The Fixed Point Set	160
IV	The Quasi-Isomorphism Theorem	162
17	The Proof in Broad Strokes	163
17.1	Pixellated Squares	163
17.2	Bad Squares	164
17.3	Errand Edges and Double Crossings	166
17.4	The Pixellation Theorem	167
17.5	The Bound Chain Lemma	168
17.6	Proof of the Quasi-Isomorphism Theorem	169
18	Proof of the Bound Chain Lemma	172
18.1	Length Two Chains	172
18.2	Length Three Chains: Case A	173
18.3	Length Three Chains: Case B	175
18.4	Length Three Chains: Case C	176
18.5	Length Four Chains: Case A	181
18.6	Length Four Chains: Case B	182
18.7	Length Four Chains: Case C	183

18.8	Length Four Chains: Case D	184
19	The Graph Reconstruction Formula	186
19.1	Main Result	186
19.2	Eliminating Most Double Crossings	188
19.3	Eliminating the last Double Crossing	189
20	The Hitset and the Intertwiner	191
20.1	The Hitset	191
20.2	The Projective Intertwiner	193
20.3	Well Definedness	195
20.4	Strategy of the Proof	196
20.5	The Intertwining Theorem on the Diagonal	197
20.6	The Hitset Theorem on the Diagonal	199
20.7	Hitset Induction	200
20.8	Changing the Fundamental Domain	202
20.9	Intertwiner Induction	204
20.10	Proof of the Geometric Claim	205
21	Correspondence of Polytopes	207
21.1	The Triple Partition	207
21.2	Proof of Statement 2	209
21.3	A Sample Result	210
21.4	Fixing Orientations	212
21.5	Edge Crossing Problems	214
22	Edge Crossings	216
22.1	The Graph Method	216
22.2	The Sample Result Revisited	218
22.3	The Plaid Method	220
22.3.1	Case 1	221
22.3.2	Case 2	221
22.3.3	Case 3	222
22.3.4	Case 4	222
22.3.5	Case 5	222
22.4	Another Edge Crossing Problem	223
22.5	Out of Bounds	223

23 Proof of the Pixellation Theorem	225
23.1 Solving Most of the Crossing Problems	225
23.2 Proof of Statement 3	225
23.3 Proof of Statement 4	226
23.4 Proof of Statement 5	228
23.4.1 Case 1	230
23.4.2 Case 2	231
23.4.3 Case 3	231
23.4.4 Case 4	232
24 Computer Assisted Techniques	233
24.1 Operations on Polytopes	233
24.2 The Calculations	235
24.3 The Computer Program	238
24.3.1 Downloading the Program	238
24.3.2 Running and Compiling	238
24.3.3 What to do first	238
24.3.4 Presets	239
24.3.5 Surveying the Data and Proofs	239
V The Distribution of Orbits	240
25 Unbounded Orbits	241
25.1 Geometric Limits	241
25.2 Approximating Irrationals	242
25.3 Arc Copying	244
25.4 The Images of the Anchors	246
25.5 The End of the Proof	248
26 Some Elementary Number Theory	250
26.1 All About Predecessors	250
26.1.1 Statement 1	251
26.1.2 Statement 2	251
26.1.3 Statements 3 and 5	252
26.1.4 Statement 4	252
26.1.5 Statement 6	253
26.1.6 Statement 7	253

26.2	Existence of the Predecessor Sequence	253
26.3	Existence of the Approximating Sequence	254
26.4	Another Identity	256
27	The Box Lemma and The Copy Theorem	258
27.1	The Weak and Strong Copying Lemmas	258
27.2	The Core Copy Lemma	259
27.3	Proof of the Box Lemma	259
27.4	Setup for the Copying Theorem	261
27.5	The Strong Case	261
27.6	The Core Case	262
28	The Weak and Strong Copy Lemmas	264
28.1	The Mass and Capacity Sequences	264
28.2	A Matching Criterion	265
28.3	Geometric Alignment	267
28.4	Alignment of the Capacity Sequences	268
28.5	Alignment of the Mass Sequences	269
29	The Core Copy Lemma	272
29.1	The Difficulty	272
29.2	Weak Horizontal Alignment	272
29.3	Geometric Alignment	273
29.4	Alignment of the Capacity Sequences	273
29.5	Calculating some of the Masses	276
29.6	Alignment the Mass Sequences	279
30	Existence of Many Large Orbits	282
30.1	The Empty Rectangle Lemma	282
30.2	The Ubiquity Lemma	283
30.3	Proof of the Main Result	284
30.4	The Continued Fraction Length	285
30.5	Circle Rotations	286
30.6	Proof of the Grid Supply Lemma	287
31	Tuned Sequences	288
31.1	Rescaling the Model	288
31.2	The Filling Property	289

31.3 Proof of Theorem 31.1	291
32 References	293

Preface

The purpose of this monograph is to study a construction, based on elementary geometry and number theory, which produces for each rational parameter (satisfying some parity conditions) a cube filled with polyhedral surfaces. When the surfaces are sliced in one direction, the resulting curves turn out to encode all the essential information about outer billiards on kites. When the surfaces are sliced in two other directions, they encode all the essential information in a 1-parameter family of the Truchet tile systems defined in [H]. I call the construction the plaid model.

The plaid model grew out of my work in [S1], where I gave an affirmative answer to the Moser-Neumann question about outer billiards: *Does there exist an outer billiards system with an unbounded orbit?* The main result of [S1] is that outer billiards has unbounded orbits relative to any irrational kite – a bilaterally symmetric convex quadrilateral which is not affinely equivalent to a lattice polygon.

The plaid model has a physical feel, with properties that seem like conservation laws, interacting particles, spacetime diagrams, and even an exclusion principle. It also has an overtly hierarchical structure, which causes it to exhibit properties such as self-similarity and scaling limits. Finally, it has an interpretation in terms of a beautiful higher dimensional polytope exchange transformation.

This monograph establishes some of the basic properties of the plaid model: the connection to outer billiards and to Truchet tilings, the connection to polytope exchange transformations, and some results about the size and distribution of the polygons in the slices of the model. I feel that the plaid model is a deep and surprising structure which blends geometry, combinatorics, number theory, and dynamics. I hope that this work points out its depth and beauty, and suggests topics for further study.

I would like to thank the National Science Foundation for their continued support, and also the Simons Foundation for an upcoming Simons Sabbatical Fellowship. I'd also like to thank Peter Doyle, Pat Hooper, Sergei Tabachnikov, and Ren Yi for a number of conversations related to the plaid model.

Introduction

Imagine starting with a square grid, as shown in Figure 0.1. One can produce interesting patterns of points on the square grid by taking some slanting lines and marking the points where they intersect the grid. The left side of Figure 0.1 shows this.

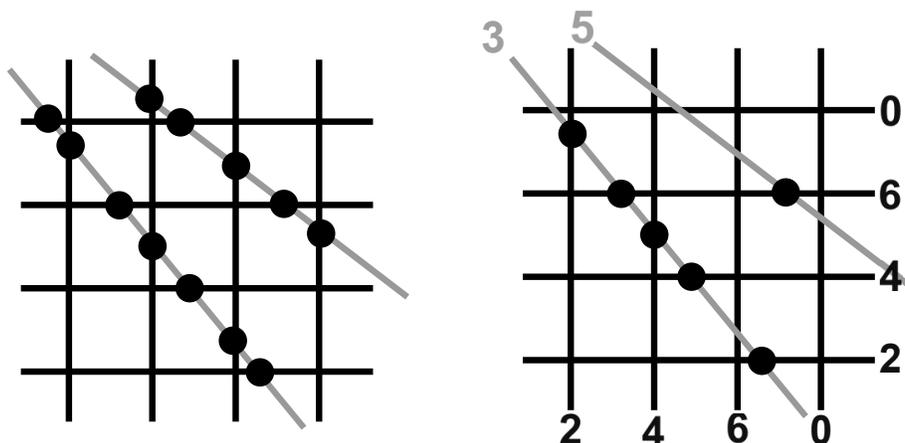


Figure 0.1: Slanting lines intersecting a grid

We could add variety to the construction by assigning “capacities” to the grid lines and “masses” to the slanting lines, and then putting intersection points only when the mass of the slanting line is less than the capacity of the grid line. The right side shows the same example as the left, but with masses and capacities added.

The plaid model starts out with a construction like this, and then we build tilings of the plane based on the arrangement of the intersection points. The tilings in turn define families of embedded polygons. Finally, we cut out various chunks (or *blocks*) of the planar tiling, stack them on top of each other, and canonically interpolate between the polygons at different levels to form embedded polyhedral surfaces.

The plaid model grew out of a theorem I proved in [S1] which I called the Hexagrid Theorem. The Hexagrid Theorem is one of the main structural tools I used to understand some properties of outer billiards on kites – enough to prove the following result.

Theorem 0.1 *Outer billiards on any irrational kite has an unbounded orbit.*

Theorem 0.1 resolved the so-called *Moser-Neumann problem*, from 1960, which asked whether an outer billiards system could ever have an unbounded orbit.

Any kite is affinely equivalent to the kite K_A with vertices

$$(-1, 0), \quad (0, 1), \quad (0, -1), \quad (A, 0). \quad (1)$$

See Figure 0.2. The kite is (ir)rational when A is (ir)rational. Figure 0.2 shows outer billiards on K_A for $A = 4/9$. Given $p_0 \in \mathbf{R}^2 - A$, we define a map $p_0 \rightarrow p_1$ by the rule that the line segment $\overline{p_0 p_1}$ is tangent to K at its midpoint, and K is on the right hand side as one walks along the segment from p_0 to p_1 . One then considers the orbit $p_0 \rightarrow p_1 \rightarrow p_2 \dots$. See [S1] for an extensive discussion of outer billiards and a long bibliography.

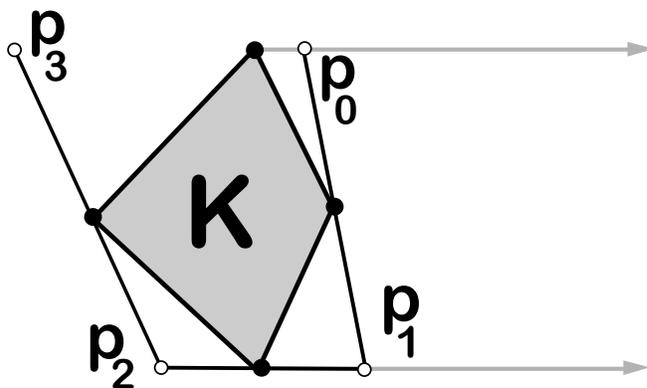


Figure 0.2: Outer billiards on the kite K .

We call an outer billiards orbit on K_A *special* if it lies in the union

$$\Xi = \{\pm 1, \pm 3, \pm 5, \dots\} \times \mathbf{R}. \quad (2)$$

Each special orbit returns infinitely often to the pair horizontal rays shown in Figure 0.2. We denote these rays by the symbol \implies .

The key step in understanding the special orbits on K_A is to associate a family of embedded lattice polygons to the orbits which encodes the symbolic dynamics of the first return map to \implies . I call these polygonal curves *the arithmetic graph*. Part 3 of the monograph has a detailed description. In [S1] The Hexagrid Theorem, mentioned above, predicts the large scale structure of the arithmetic graph based on where certain grids of lines intersect.

Some years later I discovered that the Hexagrid Theorem was just the first in a series of results which allowed this large scale structure to extend down to increasingly fine scales. When all these results are assembled into one package, the result is the plaid model. For each parameter $A = p/q$ with pq even, the plaid model construction produces a cubical array of $(p + q)^3$ unit cubes that is filled with disjoint embedded polyhedral surfaces. When these surfaces are sliced in the XY directions they produce curves which agree with the arithmetic graph associated to K_A up to one unit of error.

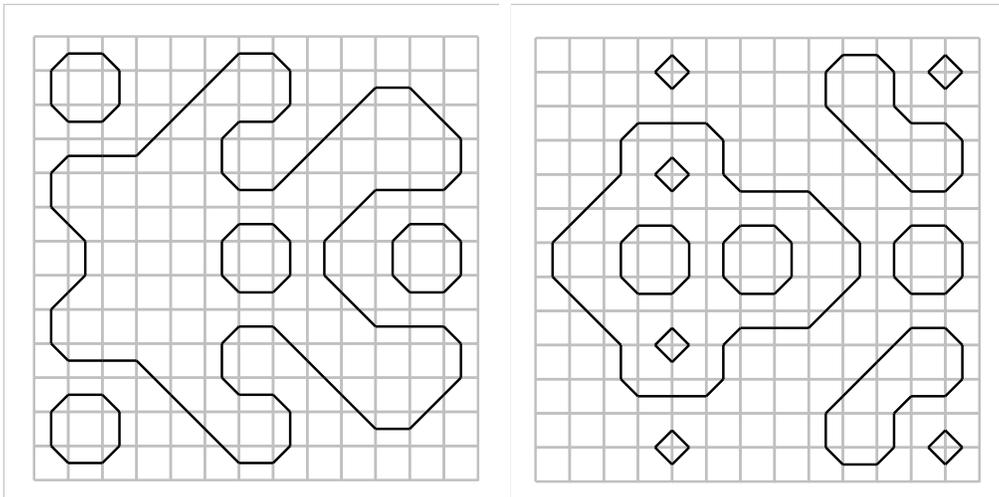


Figure 0.3: Two XY plaid model slices for the parameter $4/9$.

Figure 0.3 shows two consecutive slices of the surfaces for the parameter $4/9$. Here is a more concrete (though less powerful) way to explain the connection to outer billiards. When the polygons in Figure 0.3 are projected onto the X coordinate, each one of them agrees with the first return map of some special orbit to \implies up to a uniformly bounded error. The bound is about 4 units, and it works for all parameters.

When the surfaces in the plaid model are sliced in the other directions, namely the XZ and YZ directions, what emerges (at least for some slices) is a combinatorial pattern of curves that is combinatorially isomorphic to the curves produced by Pat Hooper's Truchet tile system [H].

Figure 0.4 gives an example. The plaid model parameter is $4/9$ and the Truchet tile parameter is $\alpha = \beta = 3/8$.

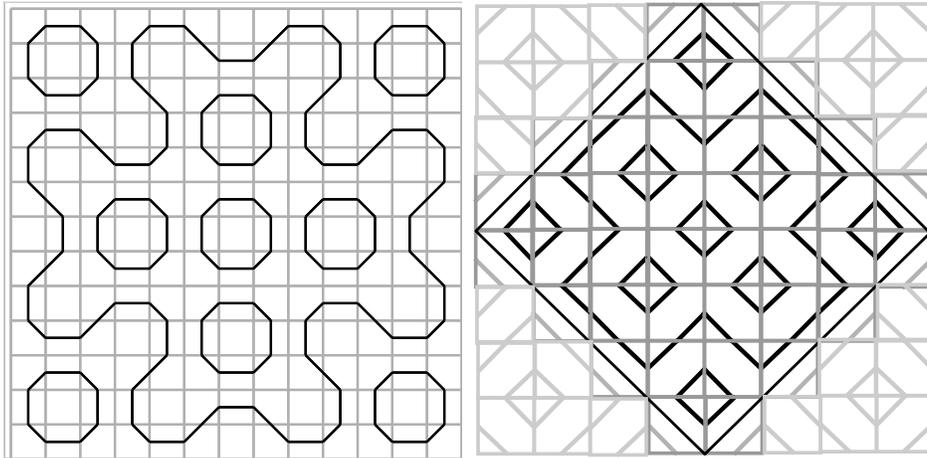


Figure 0.4: A YZ slice for $4/9$ compared to a Truchet tiling.

Thus, the 3 dimensional plaid model is a kind of marriage between the symbolic dynamics of outer billiards on kites and the Truchet tile system. The purpose of this monograph is to explore its structure. The monograph has 5 parts. Here is a detailed description of these parts.

0.1 Part 1: The Plaid Model

In Part 1, I will define the plaid model (in several ways) and study its basic properties. One thing that is not clear from any of the definitions is that the plaid model is well-defined. There will be several places where the results depend on the consistency of the definitions. I will deduce the claimed consistency results at the end of Part 2, as a consequence of the Plaid Master Picture Theorem.

After studying the basic properties of the model I will explain how one can use the hierarchical nature to get information about the large scale structure of the tilings in an algorithmic way. In particular, I will give a heuristic explanation of why the model exhibits coarse self-similarity and rescaling phenomena.

The initial construction of the plaid model produces 2 dimensional slices. I will explain how to assemble these slices into embedded polyhedral surfaces. Finally, I will establish the connection between the XZ and YZ slices of these surfaces and the Truchet tilings. The main theorem along these lines is the Truchet Comparison Theorem from §6.2.

0.2 Part 2: The Plaid Master Picture Theorem

In Part 2 I will connect the plaid model to a higher dimensional polytope exchange transformation, and I will deduce all the consistency claims made in Part 1.

An *affine polytope exchange transformation*, or *affine PET*, is given by a pair of partitions

$$X = \bigcup A_i = \bigcup B_i \quad (3)$$

of one polytope X into finitely many smaller convex polytopes $\{A_i\}$ and $\{B_i\}$ in such a way that there are affine maps T_i such that $T_i(A_i) = B_i$ for all i . We insist that the linear part of T_i is independent of i . Such a system gives rise to a global and almost everywhere defined map $T : X \rightarrow X$ defined so that $T|_{A_i} = T_i$. This map is not defined on the boundaries of the polytopes of the partition.

- When T_i is a translation for all i , the PET is called *ordinary*. The ordinary PETs are the most commonly studied.
- The affine PET is called *integral* if every vertex of every polytope in sight has integer coordinates.
- The affine PET is called *fibred* if it has an invariant codimension 1 fibration by ordinary PETs.

As above, let G denote the set of unit integer squares. For each parameter p/q , the plaid model defines a dynamical system on G . We simply follow the directed edge in each tile and move to the tile into which the edge points. When the tile is empty, we do not move at all. We call this dynamical system the p/q *plaid dynamics*. This system is similar to the *curve following dynamics* defined in [H].

Theorem 0.2 (Plaid Master Picture) *There is a 4 dimensional fibred integral affine PET X_2 with the following properties. The slice*

$$X_2(P) = X_2 \cap (\mathbf{R}^3 \times \{P\})$$

is an ordinary 3-dimensional PET. In the special case where $A = p/q$ is even rational and $P = 2p/(p+q)$, there is a locally affine map $\Phi_A : G \rightarrow X_2(P)$ which conjugates the p/q plaid dynamics on G to the PET dynamics on $X_2(P)$.

0.3 Part 3: The Arithmetic Graph

As we mentioned above, in [S1] we associate an infinite family of polygonal paths to each rational parameter $A = p/q \in (0, 1)$. These paths are embedded and have integer vertices. When pq is even, the paths are all closed embedded polygons. We call these polygons *the arithmetic graph*.

The significance of the arithmetic graph is that there is a bijection between components of the arithmetic graph and special outer billiards orbits (up to a certain trivial involution on the orbits.) The bijection has the property that the number of vertices in the component of the arithmetic graph is the same as the number of times the orbit hits the union \implies of two rays defined above. One can recover the outer billiards orbits from the graph.

Just as we have the plaid (curve following) dynamics we have the graph (curve following) dynamics for the arithmetic graph. The graph polygons are naturally oriented by the outer billiards dynamics, and we simply move from each vertex of \mathbf{Z}^2 to the next. In [S1] we proved the following result, though we stated it somewhat differently.

Theorem 0.3 (Graph Master Picture) *Let $p/q \in (0, 1)$ be any rational parameter. There is an ordinary 3 dimensional PET Y_A and a locally affine map $\Psi_A : \mathbf{Z}^2 \rightarrow Y_A$ which conjugates the p/q graph dynamics to the dynamics on Y_A . $\bigcup_A Y_A$ is dense in a 4-dimensional fibered integral affine PET.*

In this part of the monograph I will follow my preprint [S3] and put the Graph Master Picture Theorem in a wider context of higher dimensional compactifications which work for an essentially arbitrary polygonal outer billiards system. This gives a new and more conceptual proof of the Graph Master Picture Theorem.

0.4 Part 4: The Quasi Isomorphism Theorem

We say that two embedded polygons are *C-quasi-isomorphic* if they can be parametrized so that corresponding points are within C units of each other. We say that two unions Γ and Π of polygons are *C-quasi-isomorphic* if there is a bijection between the members of A and the members of Γ which pairs up C -quasi-isomorphic polygons. Finally, we say that A and B are *C-affine-quasi-isomorphic* if there is an affine transformation $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ such that Π and $T(\Gamma)$ are C -quasi-isomorphic. Here is the main result.

Theorem 0.4 (Quasi-Isomorphism) *Let $p/q \in (0, 1)$ be any rational such that pq is even. Let Π denote the plaid model at the parameter p/q . Let Γ denote the arithmetic graph associated to the special orbits of outer billiards on the kite $K_{p/q}$. Then Π and Γ are 2-affine-quasi-isomorphic.*

Figure 0.5 shows the Quasi-Isomorphism Theorem in action for the parameter $p/q = 4/9$. The black polygons are part of the plaid model and the grey polygons are part of a canonically chosen affine image of the arithmetic graph.

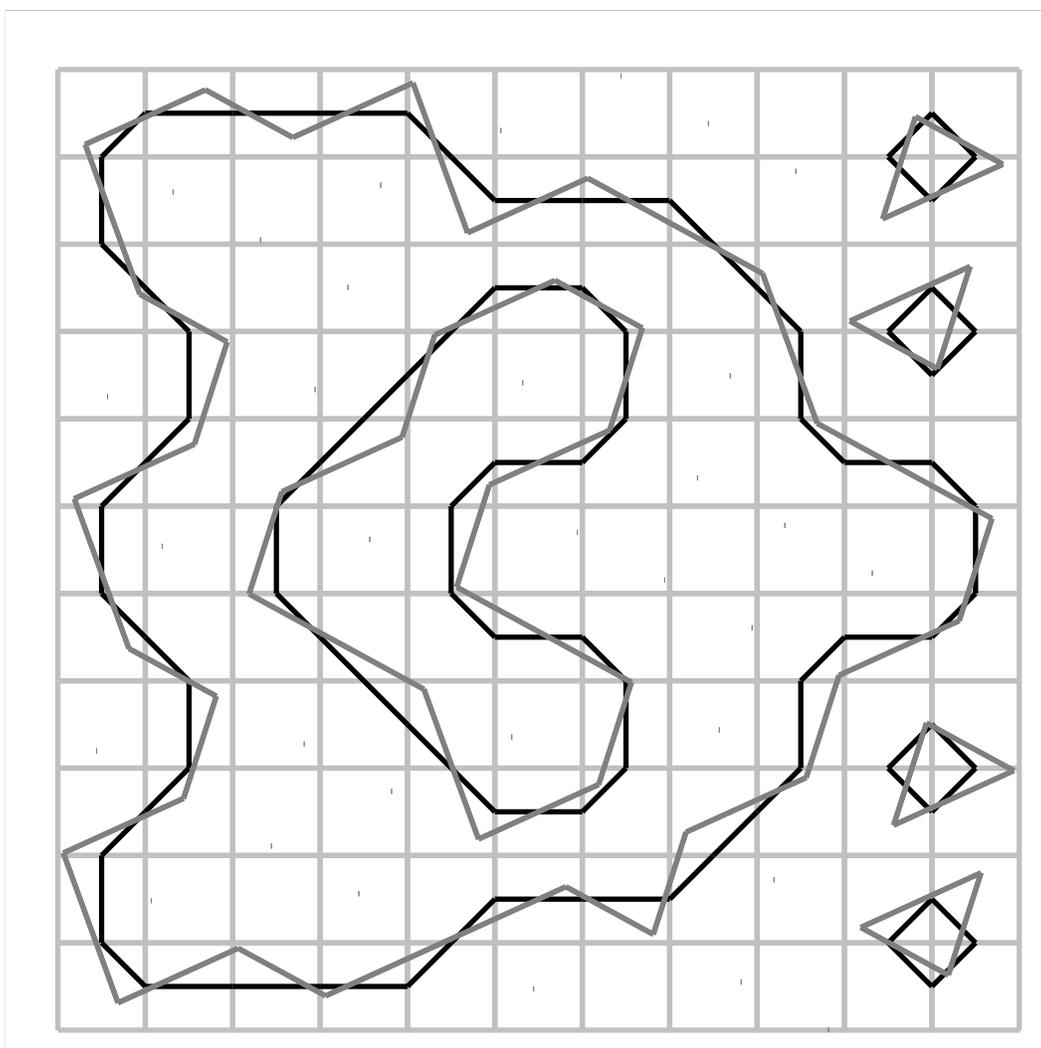


Figure 0.5: The Quasi-Isomorphism Theorem in action.

In general, the vertices of the black curves lie in the lattice $\frac{1}{2}\mathbf{Z}^2$, which has co-area $1/4$, whereas the vertices of the grey curves lie in a lattice having co-area $1 + p/q$. Even though the polygons line up in a global sense, their vertices lie on quite different lattices. This fact points to the highly nontrivial nature of quasi-isomorphism being established.

The Quasi-Isomorphism has a more dynamical formulation.

Theorem 0.5 (Orbit Equivalence) *Let X be the affine PET from the Plaid Master Picture Theorem and let Y be the affine PET from the Graph Master Picture Theorem. There is a fibered polyhedral set $Z \subset X$ which intersects every orbit on X , and a piecewise integral projective map $F : Z \rightarrow Y$ which carries the fiber Z_P to the fiber Y_A . The map F is an orbit equivalence: It sets up a bijection between the orbits in X and the orbits in Y .*

An algebraic set Z is *fibered polyhedral* if it has a codimension 1 fibration in which each fiber is a polyhedron. When X is the fibered affine PET from the Plaid Master Picture Theorem, we mean for the fibration of X to induce the fibration of Z .

The Quasi-Isomorphism Theorem is basically a consequence of the Plaid Master Picture Theorem, the Graph Master Picture Theorem, and the Orbit Equivalence Theorem, but actually the relationship between these theorems is different. We will first construct the map $F : X \rightarrow Y$. Then we will give a computer assisted proof that F has certain combinatorial properties which imply the Quasi-Isomorphism Theorem. Then we will observe at the end that we have also proved the Orbit Equivalence Theorem.

0.5 Part 5: The Distribution of Orbits

Part 5 concerns the distribution and geometry of orbits in the plaid model. Say that a set in the plane is *fat* if it is not contained in a tubular neighborhood of any straight line. A fat set, in particular, is unbounded. The Plaid Master Picture Theorem will allow us to make sense of the plaid model even at irrational parameters. With this identification, each orbit in the plaid PET corresponds to either a closed plaid polygon or an infinite polygonal path. The bulk of Part 4 is devoted to proving the following result.

Theorem 0.6 (Unbounded Orbits) *For every irrational $A \in (0, 1)$ there exists an infinite orbit in the plaid PET whose corresponding unbounded orbit is a fat path.*

The idea behind the proof of the Unbounded orbits theorem is to show that there exist orbits of large diameter in the rational case, and then to take a geometric limit. The geometric limit argument must be done carefully if we want to end up with an unbounded path that corresponds to a well-defined orbit in the plaid PET. Our analysis will reveal the importance of Diophantine considerations in the plaid model.

The Quasi-Isomorphism Theorem combines with the Unbounded Orbits Theorem to give another proof of Theorem 0.1 from [S1].

We define a *block* to be a single integer horizontal slice of the 3 dimensional plaid model. Figure 0.3 shows the picture in two different blocks. Our next result shows that all the blocks contain long orbits when the parameter is complicated.

We say that a plaid polygon is *N-fat* if it does not lie within the *N*-tubular neighborhood of any straight line.

Theorem 0.7 *Let $\{p_k/q_k\} \subset (0, 1)$ be any sequence of even rational numbers with an irrational limit. Let $\{B_k\}$ be any sequence of associated blocks. Let N be any fixed integer. Then the number of N -fat plaid polygons in B_k is greater than N provided that k is sufficiently large.*

Theorem 0.8 can be interpreted in terms of outer billiards on the kite K_A when $A = p/q$. Recall that a *special orbit* is an orbit which lies on the invariant set Ξ of horizontal lines having y intercept an odd integer. Every such orbit has points in the union \implies of two rays that we used for our first return map. Define the *essential diameter* of a special orbit O to be diameter of $O \cap \implies$.

Corollary 0.8 *Let $\{p_k/q_k\} \subset (0, 1)$ be any sequence of even rational numbers with an irrational limit. Let $\{I_k\}$ be any sequence of intervals of the form $[n\omega_k, n\omega_k + \omega_k]$. Let N be any fixed integer. Then there are more than N distinct orbits in the interval $I_k \times \{1\}$ which have essential diameter at least N provided that k is sufficiently large.*

Corollary 0.8 follows directly from Theorem 0.7 and from the Quasi-Isomorphism Theorem.

When we know more about the sequence of rationals we will be able to get detailed information. Theorem 31.1 is a more specialized result along these lines.

0.6 Companion Program

The paper comes with a companion computer program which illustrates many of the results in this paper - in particular the PET Equivalence Theorem. One can download this program from my website. The URL is

<http://www.math.brown.edu/~res/Java/PLAID2.tar>

I discovered all the results in this monograph using the program, and I have extensively checked my proofs against the output of the program. While this monograph mostly stands on its own, the reader will get much more out of it by using the program while reading. I would say that the program relates to the material here the way a cooked meal relates to a recipe.

Part I
The Plaid Model

1 Two Definitions of the Plaid Model

1.1 Basic Parameters

Throughout the monograph, we work with rational parameters $p/q \in (0, 1)$ such that pq is even. We call these parameters *even rational parameters*. Here are the main auxiliary quantities associated to these parameters.

$$\omega = p + q, \quad P = \frac{2p}{\omega}, \quad Q = \frac{2q}{\omega}. \quad (4)$$

Note that

$$\frac{P}{Q} = \frac{p}{q}, \quad P + Q = 2. \quad (5)$$

We let $\hat{\tau} \in (0, \omega)$ is the integer such that

$$2p\hat{\tau} \equiv 1 \pmod{\omega}. \quad (6)$$

Finally, we define

$$\tau = \min(\hat{\tau}, \omega - \hat{\tau}). \quad (7)$$

So, τ is the unique solution to $2p\tau \equiv \pm 1 \pmod{\omega}$ that lies in $(0, \omega/2)$.

1.2 Six Families of Lines

We consider 6 infinite families of lines.

- \mathcal{H} consists of horizontal lines having integer y -coordinate.
- \mathcal{V} consists of vertical lines having integer x -coordinate.
- \mathcal{P}_\pm is the set of lines of slope $\pm P$ having integer y -intercept.
- \mathcal{Q}_\pm is the set of lines of slope $\pm Q$ having integer y -intercept.

We call the lines in \mathcal{H} and \mathcal{V} the *grid lines*, and we call the other lines the *slanting lines*. Until the last section of this chapter, we only use the families \mathcal{V} , \mathcal{H} , \mathcal{P}_- and \mathcal{Q}_- . We set $\mathcal{P} = \mathcal{P}_-$ and $\mathcal{Q} = \mathcal{Q}_-$ in order to simplify the notation. Figure 1.1 shows these lines inside $[0, 7]^2$ for $p/q = 2/5$. In this case, $P = 4/7$ and $Q = 10/7$.

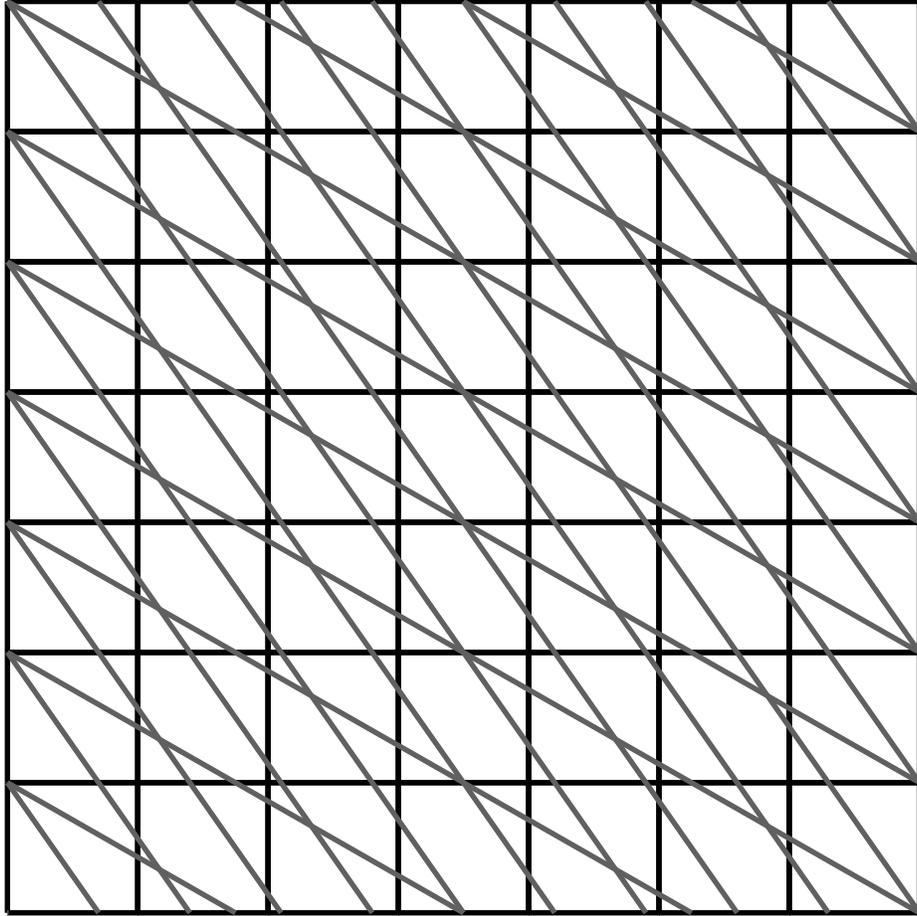


Figure 1.1: The 4 line families for $p/q = 2/5$.

Fixing the parameter p/q , we define a *block* to be the image of a unit integer square under the dilation $(x, y) \rightarrow \omega(x, y)$. For instance, $[0, \omega]^2$ is a block. The pattern of lines we have defined is precisely the same in each block.

One thing to notice about Figure 1.1 is that there are two slanting lines meeting at points on the vertical edges and the vertical midline of the block. We these points *double points*.

1.3 Capacity, Mass, and Sign

Given some integer N and some integer a , we define $(a)_{2N}$ to be the representative of $a \pmod{2N}$ in $(-N, N)$. If $a \equiv N \pmod{2N}$ we define $a_{2N} = *N$, a symbol which denotes the set $\{N, -N\}$.

We associate *signs* and *capacities* to all the lines in \mathcal{H} and \mathcal{V} as follows.

- The *signed capacity* of the line $y = y_0$ is $(4py_0)_{2\omega}$.
- The *signed capacity* of the line $x = x_0$ is $(4px_0)_{2\omega}$.
- The *signed mass* of a slanting line through $(0, y)$ is $(2py + \omega)_{2\omega}$.

Sometimes we will want to consider the signs or the absolute values of the quantities just defined. We define the *capacity* of a grid line to be the absolute value of its signed capacity. We define the *sign* of a grid line to be the sign of its capacity. Likewise we define the quantities *mass* and *sign* for the slanting lines. When the capacity or mass is 0, the sign is indeterminate. Note that the capacity is never $\pm\omega$, so it is always well defined.

Figure 1.2 shows the assignment of these quantities to the lines in the square $[0, 7]^2$. The black labels are the signed capacities and the grey labels are the signed masses.

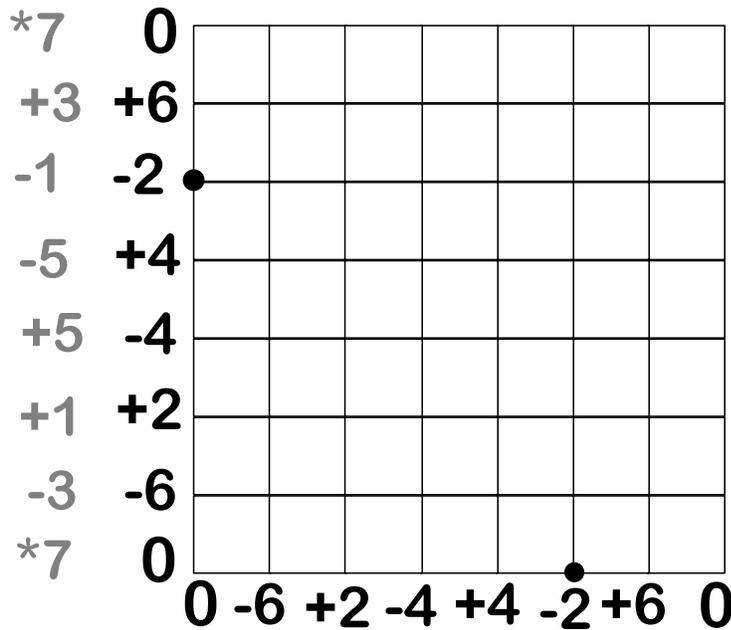


Figure 1.2: The mass and capacity labels for $p/q = 2/5$.

1.4 First Definition of the Plaid Model

We say that an *intersection point* on a grid line is the intersection of that grid line with a slanting line. Suppose that γ is a grid line and $v \in \gamma$ is the intersection of γ with a slanting line σ . We call v *light* (with respect to these two lines) if and only if

- σ and γ have the same sign.
- The mass of σ is less than the capacity of γ .

Otherwise we call v *dark*. These features of an intersection point will be called its *shade*.

We take special care when v lies on a double point.

Midline Case: Suppose that v lies on the vertical midline of a block. In this case, v lies on a horizontal grid line but not a vertical grid line. The two slanting lines through v have y -intercepts which are ω apart, so v would be reckoned light with respect to both slanting lines, or dark with respect to both slanting lines. Accordingly, we count v either as 2 light points or as 2 dark points.

Corner Case: The intersection points on the vertical edges of the block are all grid points. That is, they have integer coordinates. Such points lie on 2 slanting lines. Let v be such a point. The left and right edges of each block have capacity 0, so v is considered a dark point on the vertical line. On the other hand, if we think of v as being a point on the horizontal line that contains it, we count v as a single light point or a single dark point according to the definition above, taken with respect to either slanting line containing v .

Figure 1.3 below shows the light intersection points inside the square $[0, 7]^2$ for the parameter $2/5$. For each unit square S in this region, we have connected the center of S to the light points on ∂S .

Notice that some interesting curves seem to emerge. Notice also that there seems to be a small amount of junk, in the form of little loops, hanging off these curves. The junk occurs wherever there are two light points in a single edge. The one exception appears to be the case when the edge intersects the vertical midline (right in the center of the picture) but in this case the one light point on each relevant horizontal edge is counted twice.

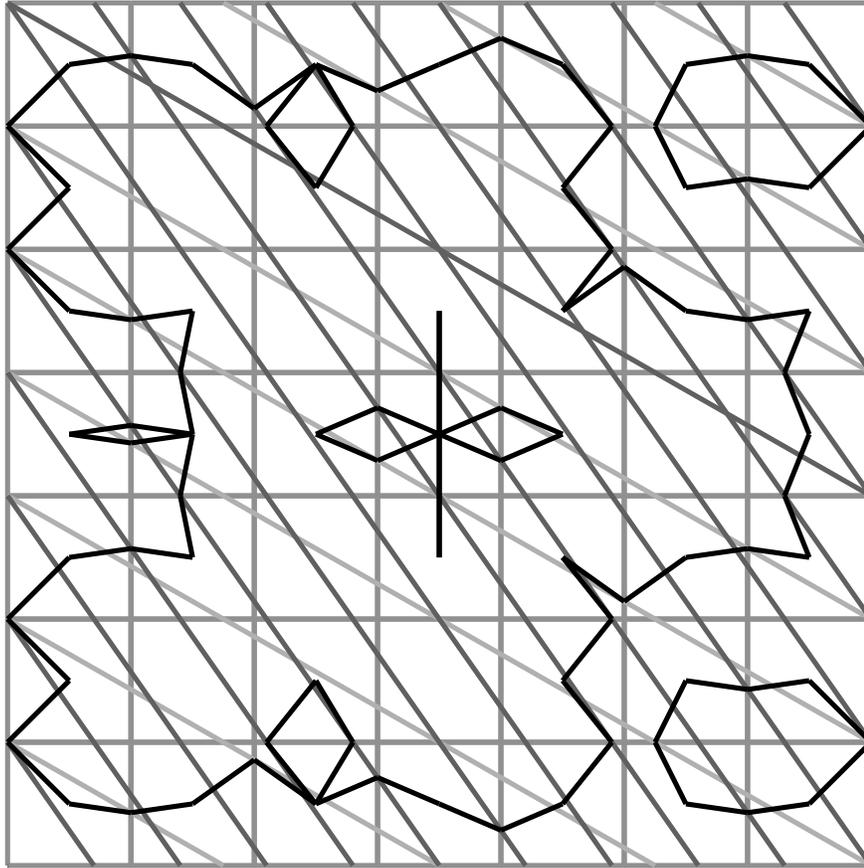


Figure 1.3: The light points in $[0, 7]^2$ with respect to $p/q = 2/5$.

We say that a *unit integer square* is a unit square whose vertices have integer coordinates. Such squares are bounded by grid lines. We call the edges of such a square *unit segments*. Such segments, of course, either lie on \mathcal{H} lines or on \mathcal{V} lines.

We say that a unit segment is *good* if it contains exactly one light point. The points are reckoned with respect to the grid containing the edge. Thus, for the boundary cases considered in the previous section, the point might be light with respect to the horizontal edge containing it, but it would always be dark with respect to the vertical edge containing it.

We say that a unit square is *coherent* if it contains either 0 or 2 good segments. We say that the plaid model is *coherent* at if all squares are coherent for all parameters. Here is the fundamental theorem concerning the plaid model.

Theorem 1.1 *The plaid model is coherent for all parameters.*

Theorem 1.1 is an immediate consequence of the Plaid Master Picture Theorem, which we prove in Part 2 of the monograph.

Theorem 1.1 allows us to create a union of embedded polygons. In each unit square we draw the line segment which connects the center of the square with the centers of its good edges. The squares with no good edges simply remain empty. We call these polygons the *plaid polygons*. Figure 1.4 shows the plaid polygons contained in $[0, 7]^2$ for the parameter $2/5$. Of course, these plaid polygons fill the entire plane, and we are just showing a small part of the picture.

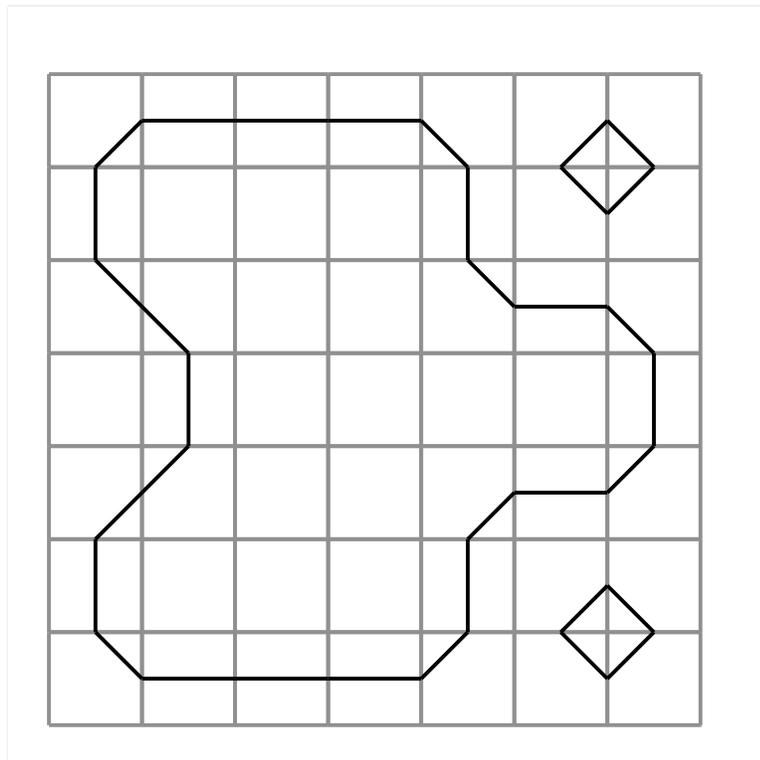


Figure 1.4: The plaid polygons inside $[0, 7]^2$ for $p/q = 2/5$.

1.5 Second Definition of the Plaid Model

Now we give another definition of the plaid model which only uses the signs of the lines. This definition uses all 6 families of lines defined above.

Horizontal Intersection Points: Reflection in any horizontal grid line preserves the grid and maps the slanting lines of negative slope to the slanting lines of positive slope. Therefore, every intersection point on a horizontal line is contained in 2 slanting lines. The two slanting lines containing a horizontal intersection point have the same slope up to absolute value. Let $\epsilon_0, \epsilon_+, \epsilon_- \in \{-1, 1\}$ respectively denote the sign of the horizontal line through v , the positive slanting line through v , and the negative slanting line through v . We call v *light* if and only if

$$\epsilon_0 = \epsilon_- = \epsilon_+. \quad (8)$$

Notice that this definition does not mention the masses or capacities of any of the lines involved.

Vertical Intersection Points: Thanks to the fact that $P + Q = 2$, every vertical intersection point v lies on a positive slanting line and a negative slanting line. The sum of the absolute values of the slopes of these slanting lines is 2. Making the same definitions as in the horizontal case, we say that v is light if and only if

$$\epsilon_\infty = \epsilon_- = -\epsilon_+. \quad (9)$$

Here ϵ_∞ is the sign of the vertical line through v . Note the $(-)$ sign in front of ϵ_+ in this definition.

Special Cases: Suppose that v is an intersection point on the vertical midline of a block. In this case v lies on 4 slanting lines. The sign of the slanting line of slope $\pm P$ coincides with the sign of the slanting line of slope $\mp Q$, and so v would be reckoned light or dark whether we used the P -lines or the Q -lines in the definition. So, again, we count v either as 2 light points or as 2 dark points.

When v lies on the vertical edge of a block, and we consider v as a point on a vertical edge, we have $\epsilon_\infty = 0$ and so Equation 9 never holds. Thus, v is reckoned as dark. If v is considered as a point on a horizontal segment, then we use Equation 8 to determine whether v is light or dark. In this case we have $\epsilon_+ = \epsilon_-$ automatically.

Lemma 1.2 *The two definitions of the plaid model coincide.*

Proof: To prove this result, we define the signed mass of a positive slanting line just as in the negative case. Let $v = (x, y)$ be a horizontal intersection point. Suppose that v lies on slanting lines of slope $\pm P$. The case when v lies on slanting lines of slope $\pm Q$ is similar. The signed capacity of the horizontal line through v is $\kappa = (4py)_{2\omega}$. By symmetry it suffices to consider the case when $\kappa > 0$. The signed capacities of the slanting lines through v are given by

$$\mu_{\pm} = (2py \mp 4p^2x)_{2\omega} \quad (10)$$

Note that $\mu_+ + \mu_- \equiv \kappa \pmod{2\omega}$.

If v is light according to the first definition, then $0 < \mu_- < \kappa < \omega$. But then μ_+ and $\kappa - \mu_-$ are both lie in $(-\omega, \omega)$ and are congruent mod 2ω . This forces

$$\mu_+ + \mu_- = \kappa. \quad (11)$$

But then all the signs agree. Hence v is light according to the second definition. Conversely, if v is light according to the second definition, then $\mu_{\pm} \in (0, \omega)$. Again, this forces $\mu_+ + \mu_- = \kappa$. But then $0 < \mu_- < \kappa$ and v is light according to the first definition.

Now consider the vertical case. We will suppose that $v = (x, y)$ is an intersection point contained on a slanting line of slope $-P$ and a slanting line of slope Q . The other case is similar. We define κ as above. Again, we will consider the case when $\kappa > 0$. The other case has a similar treatment. The quantity μ_- computes the signed mass of v relative to the P -line and we define

$$\lambda_+ = (2py - 4pqx)_{2\omega}. \quad (12)$$

The sign of the Q line is the sign of λ_+ . We have that $\mu_- - \lambda_+ \equiv \kappa \pmod{2\omega}$.

If v is light according to the first definition, then $0 < \mu_- < \kappa < \omega$. Since $\lambda_+ \in (-\omega, \omega)$ and $\lambda_+ \equiv \mu_- - \kappa$ we must have $\lambda_+ < 0$. Hence v is light according to the second definition. Conversely, if v is light according to the second definition, then we have $\mu_-, \kappa \in (0, \omega)$ and $\lambda_+ \in (-\omega, 0)$. The congruence condition above forces $\kappa = \mu_- - \lambda_+$. But then $\mu_- < \kappa$ and v is light according to the first definition. ♠

1.6 The Directed Version

Now we explain an enhanced version of the plaid model, in which the light points have transverse directions.

Definition of Direction: We have $\omega = p + q$. Also define $\omega' = q - p$. Given a slanting line L , let y be the y intercept of L . Consider the quantity

$$\delta(L) = (\omega'y)_{2\omega}. \quad (13)$$

We leave L undirected if $\delta(L) = 0$. We direct L downward if $\delta(L) < 0$ and upward if $\delta(L) > 0$. Let Λ denote a grid line and suppose that $p = L \cap \Lambda$ is a light point. We attach an arrow to p which points perpendicular to Λ and makes a positive dot product with the oriented version of L . We call such a point a *directed light point*. Each light point lies on 2 directed lines and so there is a question of consistency.

Lemma 1.3 *The two definitions of the direction are consistent.*

Proof: Let $\sigma(y)$ denote the sign of the slanting lines through $(0, y)$. We have $\sigma(y) = 1$ if and only if $(2py + \omega)_{2\omega} > 0$. This is true if and only if $(2py)_{2\omega} < 0$. Let $\delta(y)$ denote the sign of $(\omega'y)_{2\omega}$. Using the fact that $2p = \omega - \omega'$ we see that $\delta(y) = \sigma(y)$ if and only if y is odd. In short, we can deduce the direction on a slanting line from its sign and the parity of its y -intercept.

Consider a horizontal light point $\zeta = (x, m)$ for some $m \in \mathbf{Z}$. From the second definition of the plaid model, the two slanting lines through ζ have the same sign. The y -intercepts are integers equidistant from m . Hence, these y -intercepts have the same parity. But then both slanting lines point up or down by definition.

Consider a vertical light point $\zeta = (m, y)$ for some $m \in \mathbf{Z}$. One of the slanting lines through ζ has slope $\pm P$ and the other one has slope $\mp Q$. In any case, the difference between the y -intercepts is $(P + Q)m = 2m$. Hence the y -intercepts have the same parity. But the slanting lines through these points have opposite sign, thanks to the second definition of the plaid model. Hence one of the slanting lines points up and the other points down. Hence, they both point left or they both point right. In either case, they assign the same direction to the light point. ♠

Example: Figure 1.5 shows an example of two light points on the tile with center $(11/2, 1/2)$ for the parameter $2/9$. The corresponding slanting lines have y -intercept 2 and 3. Since $(14)_{22} = -8 < 0$ and $(21)_{22} = -1 < 0$, both lines point downward.

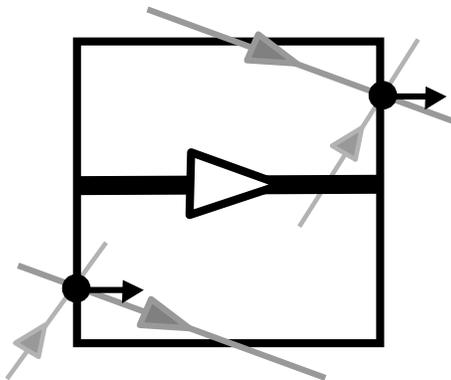


Figure 1.5: Two directed light points associated to a tile.

Tile Consistency: There is a second kind of consistency we notice with the directions. Call a light point *relevant* if it is the only light point on its unit integer edge. Theorem 1.1 says that there are either 0 or 2 relevant light points associated to each unit square. The two shown in Figure 1.5 are both relevant. We call the assignment of directions *consistent* at a tile if the same number of relevant directed light points point into the tile as point out of it. The example shown in Figure 1.5 is consistent at the tile. We will prove, as a consequence of Theorem 7.5, that the assignment of directions is consistent at every tile and for every parameter. This allows us to canonically orient each plaid polygon.

2 Basic Properties of the Model

2.1 A Characterization of the Masses and Capacities

In this section we reformulate the rules we discussed in §1.3. We fix an even rational p/q and let ω and $\widehat{\tau}$ be as in §1.1

Lemma 2.1 *For $k = 0, \dots, (\omega - 1)/2$, the lines of capacity $2k$ have the form*

$$x = k\widehat{\tau}, \quad x = \omega - k\widehat{\tau}, \quad y = k\widehat{\tau}, \quad y = \omega - k\widehat{\tau}.$$

For $k = 1, 3, \dots, (\omega - 1)$, the the lines of mass k have y -intercepts

$$(0, k\widehat{\tau}), \quad (0, \omega - k\widehat{\tau}).$$

These equations are taken mod ω .

Proof: We will deal with the line $y = k\widehat{\tau}$. When $y = \omega - k\widehat{\tau}$ the computation is similar. Note that $2p\widehat{\tau} \equiv 1 + \omega \pmod{2\omega}$ because $2p\widehat{\tau}$ is even. This congruence gives us

$$4py = (2p\widehat{\tau})(2k) \equiv 2k \pmod{2\omega}.$$

Hence

$$(4py)_{2\omega} = 2k.$$

The proof for the vertical grid lines works the same way.

We deal with the \mathcal{P} lines and \mathcal{Q} lines at the same time. We will deal with the case when the y -intercept $k\widehat{\tau}$. The other case is similar. We now are assuming that k is odd. We compute

$$2py + \omega = (2p\widehat{\tau})k \equiv k + k\omega + \omega \equiv k \pmod{2\omega}.$$

Hence the given line has mass k . ♠

The close connection between the plaid model and circle rotations suggests that there ought to be a lot of renormalization going on in the model. We will not pursue this here, though we will exploit the circle rotation property at various times in Part 4 of the monograph.

2.2 Symmetries

The Symmetry Lattice We fix some even rational parameter p/q . Let $\omega = p + q$ as above. Let $L \subset \mathbf{Z}^2$ denote the lattice generated by the two vectors

$$(\omega^2, 0), \quad (0, \omega). \quad (14)$$

We call L the *symmetry lattice*.

Blocks: We have already defined the notion of a block. We repeat the definition here for emphasis. We define the square $[0, \omega]^2$ to be the *first block*. The pictures above always show the first block. In general, we define a *block* to be a set of the form $B_0 + \ell$, where B_0 is the first block and $\ell \in L$. With this definition, the lattice L permutes the blocks. We define the *fundamental blocks* to be $B_0, \dots, B_{\omega-1}$, where B_0 is the first block and

$$B_k = B_0 + (k\omega, 0). \quad (15)$$

The union of the fundamental blocks is a fundamental domain for the action of L . We call this union the *fundamental domain*. Note that plaid polygons never intersect the boundary of a block, because the boundary lines have capacity 0.

Translation Symmetry: The assignment of signed masses and signed capacities to the integer points on the y -axis is invariant under translation by $(0, \omega)$. Likewise the assignment of signed capacities to the integer points along the x -axis is invariant under translation by $(\omega, 0)$. Finally, if we translate by the vector $(0, \omega^2)$, each slanting line is mapped to another slanting line of the same type whose y -intercept has been translated by either $P\omega^2$ or $Q\omega^2$, both of which are even multiples of ω . Hence, the plaid model is invariant under the symmetry lattice L . The picture in any block is translation equivalent to the picture in a fundamental block.

Rotational Symmetry: Reflection in the origin preserves all the masses and capacities and reverses all the signs, and hence is a symmetry of the plaid model. Combining this with the translation symmetry, we see that reflection in the center of the fundamental domain is also a symmetry of the plaid model. This center is the center of the block $B_{\omega-1}/2$. Reflection in the midpoint of a vertical side of a block is also a symmetry of the plaid model.

Reflection Symmetry: From the second definition of the plaid model, we see that reflection in the coordinate axes are symmetries of the plaid model. Such reflections either preserve or reverse the signed capacities, and permute the special families of lines. Combining this reflection symmetry with the translation symmetry, we see that reflection in the horizontal midline of a block is a symmetry of the plaid model. This explains the bilateral symmetry one sees in Figures 1.3 and 1.4.

Combining the reflection symmetry and the rotation symmetry, we see that the block $B_{(\omega+1)/2}$ always has 4-fold dihedral symmetry. Figure 2.1 shows B_{10} with respect to the parameter $5/14$.

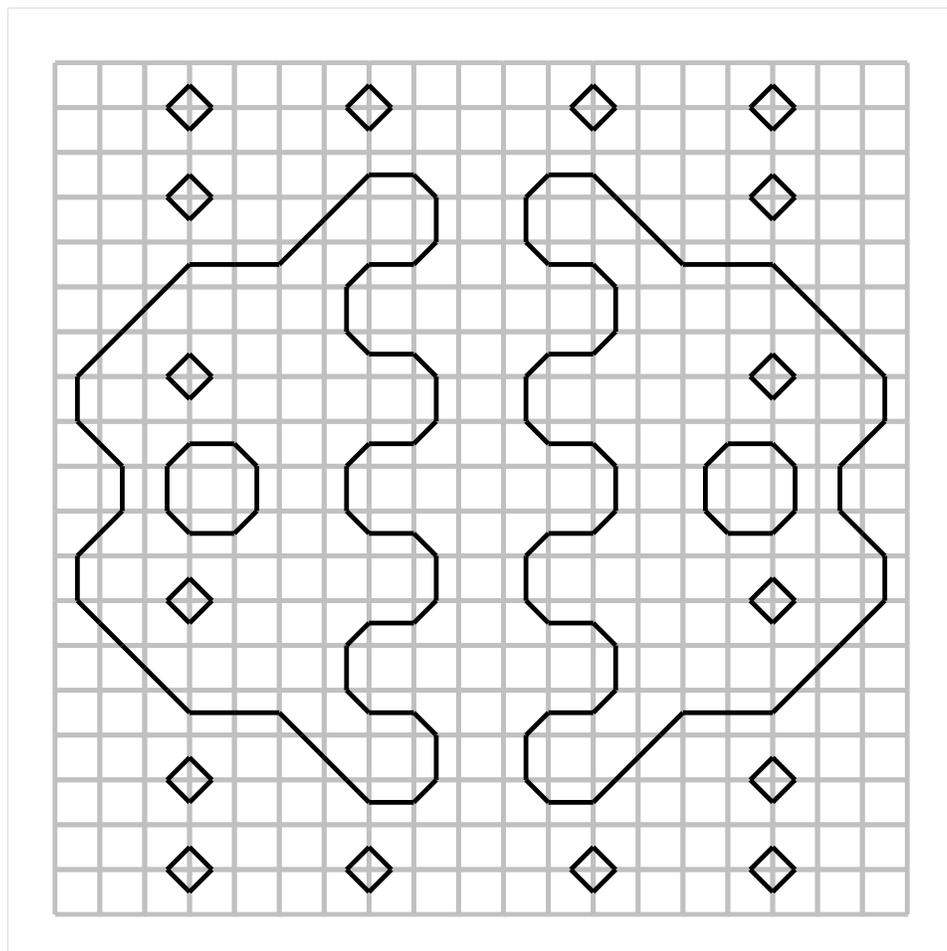


Figure 2.1: B_{10} for the parameter $5/14$.

2.3 Symmetry and Direction

How we discuss how the symmetry lattice interacts with the directions attached to the plaid model in §1.6.

Lemma 2.2 *We have the following symmetry:*

1. *Translation by $(\omega, 0)$ reverses the directions.*
2. *$(\omega^2, 0)$ respects the directions.*
3. *Rotation in the origin respects the directions.*
4. *Reflection in the coordinate axes respect the directions.*
5. *Reflection in the horizontal midline of a block reverses the directions.*

Proof: In the case of $(0, \omega)$ the key observation is that the vector $(q - p)\omega$ is congruent to $\omega \pmod{2\omega}$. This fact implies that translation by $(0, \omega)$ reverses the assigned directions. This proves Statement 1.

Consider the vector $(\omega^2, 0)$. Consider a slanting line L_0 of slope $-P$ and y intercept y_0 . Let $L_1 = L_0 + (\omega^2, 0)$. The y -intercept of L_1 is

$$y_1 = y_0 + \omega^2 P = y_0 + 2p\omega.$$

But then

$$(q - p)(y_1 - y_0) \equiv 0 \pmod{2\omega}.$$

This proves that L_0 and L_1 have the same direction. The same argument works for the other slanting lines. This proves Statement 2.

Reflection in the origin preserves the orientation of every slanting line and permutes the light points. This proves Statement 3.

Reflection in the x -axis interchanges the positive slanting lines with the negative slanting lines but preserves their orientations. This proves Statement 4 for reflection in the x -axis. Statement 4 for the y -axis follows from Statement 3 and from Statement 4 for the x -axis.

Statements 5 is a consequence of Statements 1 and 4. ♠

When we consider the directed model modulo translation symmetry, we have to consider 2ω fundamental blocks rather than ω fundamental blocks. We will discuss this below in more detail.

2.4 The Number of Intersection Points

The purpose of this section is to prove the following result.

Lemma 2.3 *Each unit segment contains 2 intersection points.*

Proof: Let e be a vertical edge. Let L be the vertical line through e . There is some $\alpha \in [0, 1]$ such that the intersection points of type P along L have the form $n + \alpha$, where $n \in \mathbf{Z}$. The same goes for the points of type Q . Hence, there are exactly two of them in e . (This works even when $\alpha = 0$.)

Now let e be a horizontal edge. If we forget about whether the intersection points are light or dark, the whole picture is symmetric under translation by $(0, 1)$ and also $(p + q, 0)$. So, we can assume that e lies on the south boundary the first block. The intersection points of type P have the form $(n/P, 0)$ where $n \in \mathbf{Z}$ and the intersection points of type Q have the form $(n/Q, 0)$.

Case 1: If e is the central edge, then e contains the intersection points

$$(p/P, 0) = (q/Q, 0) = ((p + q)/2, 0).$$

This common point is counted twice, by convention.

Case 2: If e is the westernmost edge, then e contains the two intersection points $(0, 0)$ and $(1/Q, 0)$. If e is the easternmost edge, then e contains $(p + q, 0)$ and $(p + q, 0) - (1/Q, 0)$.

Case 3: If e is not one of the edges above, then neither the boundary nor the midpoint of e contains an intersection point. Since $1/Q \in (1, 2)$ we know that e contains at least 1 point of type Q and at most 2 of them. We will show that if e does not contain a second point of type Q then e contains a point of type P . Let $(k_1/Q, 0)$ be the point of type Q that e does contain. We must have $k_1 \in (Qm + Q - 1, Qm + 1)$, for otherwise we could add or subtract 1 from k_1 and produce another intersection point of type Q in e . We claim that there is some point of type P inside e . We seek a point $k_2 \in (Pm, Pm + P)$. This time we have

$$k_1 + k_2 \in (2m + Q - 1, 2m + P + 1) = (2m + 1 - P, 2m + 1 + P),$$

The value $k_2 = (2m + 1) - k_1$ does the job. ♠

2.5 Capacity and Mass

Now we come to a more subtle result which suggests the hierarchical nature of the plaid model. The lines of small capacity have very few light points, so they predict something about the large scale geometry of the loops in the model. As we add more lines of higher capacity, the picture of the loops fills in at finer scales. In the next chapter we take this up in detail.

Theorem 2.4 *Let B be any block. For each even $k \in [0, p + q]$ there are 2 lines in \mathcal{H} and 2 lines in \mathcal{V} which have capacity k and intersect B . Each such line carries k light points in B .*

Lemma 2.5 *Statement 1 of Theorem 2.4 is true.*

Proof: Recall that $\omega = p + q$. Given the periodicity of the capacity labels, it suffices to prove this for the first block. We will prove the result for \mathcal{H} . The result for \mathcal{V} has virtually the same proof. We are simply trying to show that there are exactly 2 integer values of y in $[0, \omega]$ such that $4py = \pm k \pmod{2\omega}$. Writing $k = 2h$, we see that this equation is equivalent to $2py = \pm h \pmod{\omega}$. This has 2 solutions mod ω because $2p$ is relatively prime to ω . ♠

Lemma 2.6 *Theorem 2.4 holds in the vertical case.*

Proof: Let L be a vertical line of capacity k . The case $k = 0$ is trivial, so we assume $k > 0$. In this case, no point of type P coincides with a point of type Q. We will show that there are $k/2$ light points of type P in $L \cap B$. By reflection symmetry, the same goes for the points of type Q.

We will suppose that L has positive sign. The number of light points of mass at most k equals the number of equivalence classes mod ω of integers y such that $0 < (2py + \omega)_{2\omega} < k$. We can achieve all the values $1, 3, \dots, (k - 1)$. This gives $k/2$ light points of type P. At the same time, if

$$(2py_1 + \omega)_{2\omega} = (2py_2 + \omega)_{2\omega},$$

then $y_1 \equiv y_2 \pmod{\omega}$. Thus, there are exactly $k/2$ light points of type P on L , as claimed. ♠

Lemma 2.7 *Theorem 2.4 holds in the horizontal case.*

Proof: It suffices to prove this when B is one of the fundamental blocks, and it suffices to consider horizontal lines of positive capacity. Let L be the horizontal line of signed capacity $+k$. We fix some odd $\ell \in \{1, 3, \dots, k-1\}$.

Let S_Q (respectively \widehat{S}_P) denote the set $m \in \mathbf{Z}$ such that the line of slope $-Q$ (respectively $+P$) through $(0, m)$ intersects L . The light points on L of mass ℓ correspond to integers $m \in S_Q \cup \widehat{S}_P$ having

$$2pm + \omega \equiv \ell \pmod{2\omega}. \quad (16)$$

We just have to show that there are exactly two such points on L having this description.

We claim $\widehat{S}_P \cup S_Q$ contains exactly one number in each equivalence class mod 2ω except for the two extreme values in the set (i.e. largest and smallest), which are congruent to each other. Consider first the case when B is the first block. In this case, \widehat{S}_P consists of $2p+1$ consecutive integers starting at the left endpoint of L and going down, while S_Q consists of $2q+1$ consecutive integers starting at the left endpoint of L and going up. So, in this case, the union is a run of $2\omega+1$ consecutive integers. Our claim is true in this case. When we replace the first block B by the k th block B' , the set \widehat{S}_P moves down by $2pk$ units and the set S_Q moves up by $2qk$ units. Since $2p+2q=2\omega$, the claim still holds.

Now we know that $\widehat{S}_P \cup S_Q$ contains numbers in every congruence class. Since 2 and p are both relatively prime to ω , Equation 16 has exactly 2 solutions mod 2ω , and these two solutions differ by an odd multiple of ω . The two light points on L coincide only if (after relabeling) $m \in \widehat{S}_P$ and $m' \in S_Q$. But in this case, since $P+Q = P - (-Q) = 2$, the two light points must lie on the vertical midline of the block. We have agreed to count these points twice. So, in all cases, we have at least 2 light points of mass ℓ on L .

Given the structure of $\widehat{S}_P \cup S_Q$, it can happen that there are three solutions to Equation 16 in the set. In this case, two of the solutions must be the extreme points, and the corresponding slanting lines through them meet at the vertical edge of the block. In this case, we have agreed already to count this light point just once. ♠

3 Using the Model

In this chapter we will explore some properties of the plaid model which assume Theorem 1.1. Our main point here is to feature the hierarchical nature of the plaid model and explain how this gives us information about the nature of the orbits.

3.1 The Big Polygon

Let p/q be an even rational parameter and let $\omega = p + q$. Here we will show that there is always a polygon $\Gamma_{p/q}$ having diameter at least $q/2$ associated to the rational parameter p/q . This result will be useful when we prove the Unbounded Orbits Theorem.

Theorem 3.1 *Let B denote the first block with respect to p/q . Then there exists a plaid polygon Γ whose projection onto the x -axis has diameter at least $q/2$. Moreover, Γ has bilateral symmetry with respect to reflection in the horizontal midline of B .*

Proof: Let L be the horizontal line of capacity 2 and positive sign which intersects B . Let $z_1 = (0, y) \in L$. By Lemma 2.1, we know that z_1 is a light point of mass 1. Let $z_2 = (\omega^2/2q, y)$. We compute that z_2 is another light point on L . Since L has capacity 2, Theorem 3.1 says that these are the only two light points on $L \cap B$. The plaid which crosses the unit horizontal segment containing z_1 must also cross the unit horizontal segment containing z_2 because it has to intersect $L \cap B$ twice. This shows that the projection of Γ onto the x -axis has diameter at least $\omega^2/(2q) - 1$. Finally, we note that $\omega^2/(2q) - 1 > q/2$.

Let Γ' denote the reflection of Γ in the horizontal midline of B . We want to show that $\Gamma' = \Gamma$. Let V_1 and V_2 denote the two vertical lines of B having capacity 2. These lines are symmetrically placed with respect to the vertical midline of B . Hence, one of the two lines, say V_1 , lies less than $\omega/2$ units away from the y -axis. Since $\omega/2 < \omega^2/(2q)$, the point z_2 is separated from the y -axis by V_1 . Hence both Γ and Γ' intersect V_1 . Since there can be at most 1 plaid polygon which intersects $V_1 \cap B$, we must have $\Gamma = \Gamma'$. ♠

We call Γ *the big polygon*. Figure 3.1 shows some examples. The lines of capacity 2 are also shown in the figure.

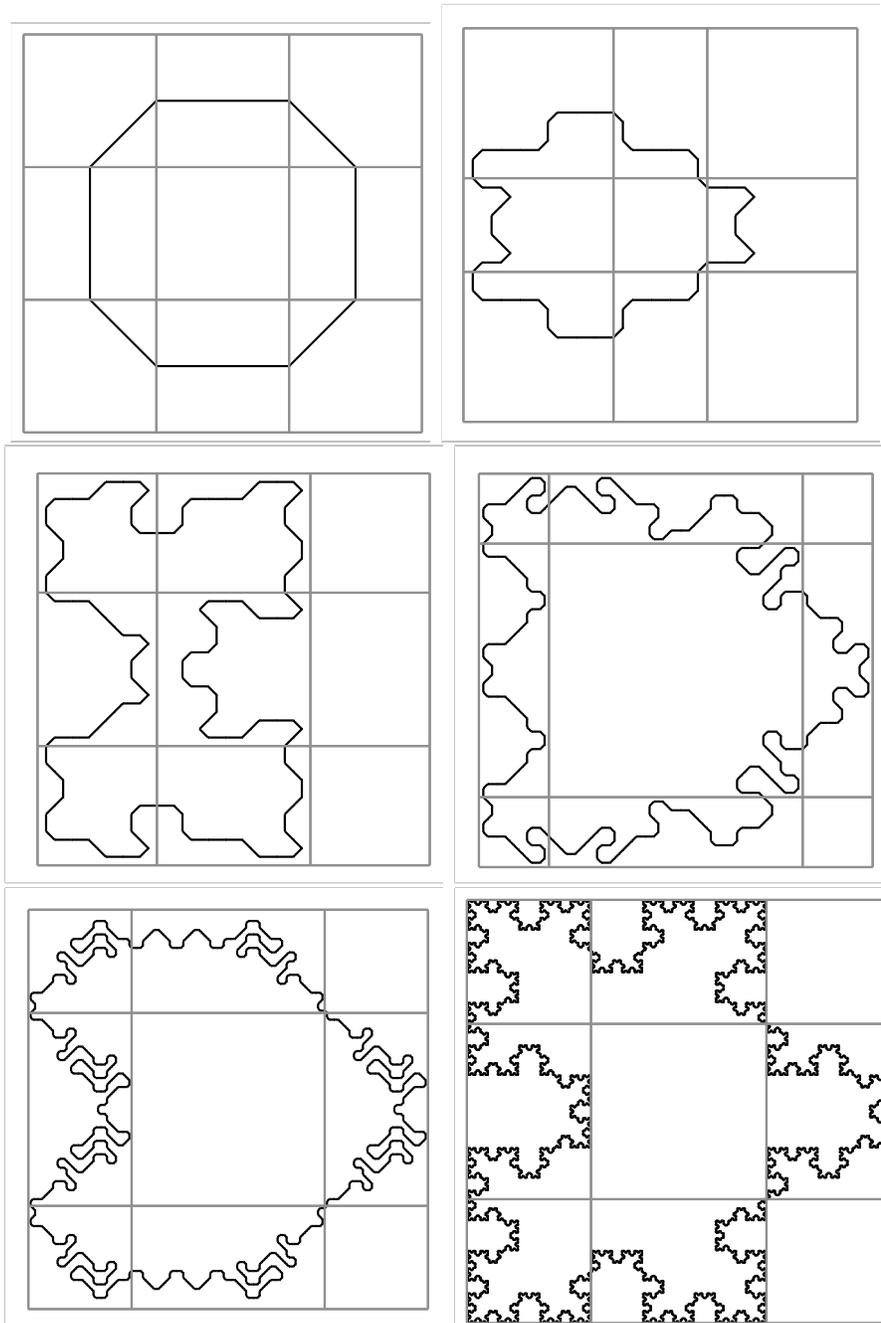


Figure 3.1: The picture for $1/2$, $4/17$, $5/18$, $14/31$, $29/60$, $169/408$.

3.2 Hierarchical Information

In establishing the existence of the big polygon, we used a very small amount of the plaid model, just the lowest capacity lines. In the next few sections we explain how to get increasingly fine scale information using more of the model. We will work with the directed plaid model because it gives more information.

We say that a *directed point* on the boundary of a rectangle is a point on this boundary equipped with an arrow which either points out of the rectangle or into it. We say that a *decorated rectangle* is a rectangle with finitely many directed points on its boundary.

Given an decorated rectangle, we say that a *connection pattern* associated to the rectangle is a finite union of connectors which uses all the points. So, if there are $2n$ directed points, there are n connectors. (There might not be any connection patterns at all.) Figure 3.2 shows several examples. Notice that in the last two examples, the same set of directed points admits two different connection patterns.

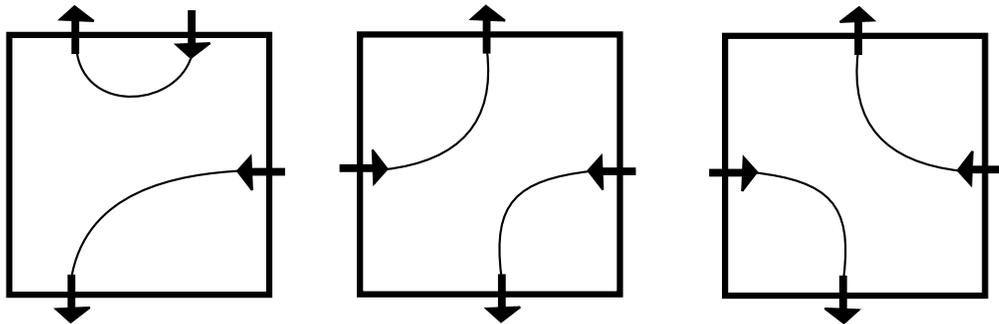


Figure 3.2: Connection patterns.

We call a decorated rectangle *unambiguous* if there is at most 1 connection pattern associated to it, in a combinatorial sense. Otherwise we call the rectangle *ambiguous*.

Suppose now that we have some even rational parameter p , some block B , and some integer $K \geq 0$. When we consider the partial grid Γ_K we get a grid of $(K+1)^2$ rectangles. By considering the light points on each rectangle, we see that each rectangle in the grid is naturally an directed rectangle. We call the triple $(p/q, B, K)$ *unambiguous* if each of the associated rectangles is unambiguous.

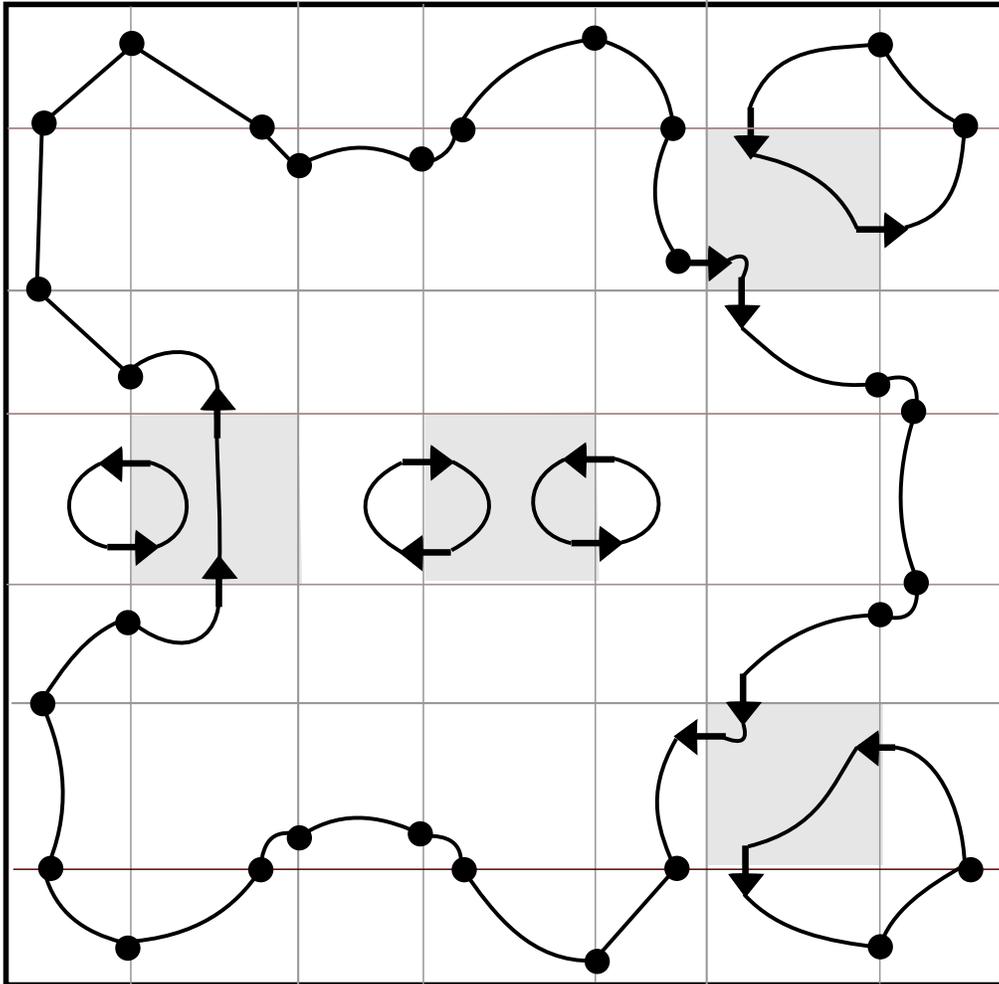


Figure 3.4: The connection pattern in the first block for $5/12$ when $K = 3$.

Figure 3.4 is an elaboration of Figure 3.3. First of all, we now can see that there are 5 additional large plaid polygons in the first block. Second of all, we can see the very largest of these polygons somewhat more precisely. The largest polygon in Figure 3.4 now gives a pretty good approximation to the actual picture in the first block for the parameter $5/12$. This is shown on the left side of Figure 3.5.

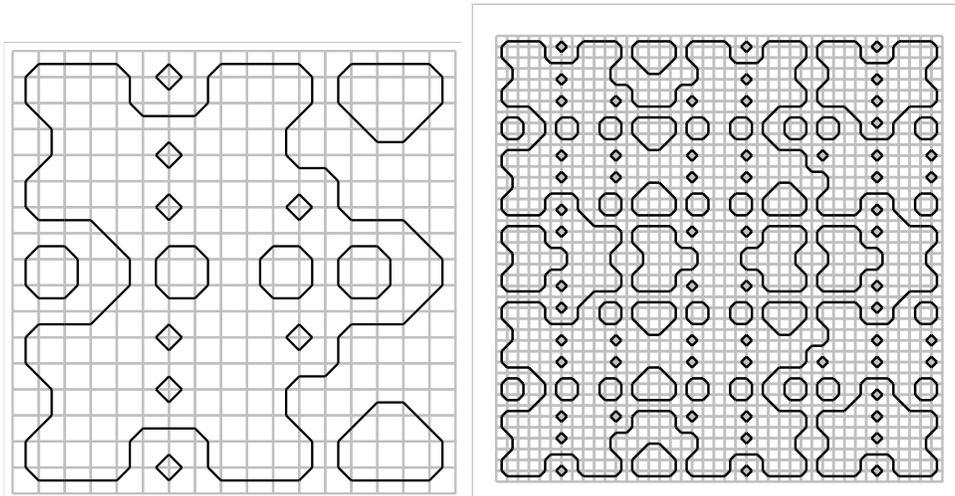


Figure 3.5: The first block for the parameter $5/12$ and $12/29$.

This process of gleaning large scale information from low capacity lines suggests the mechanism by which the large scale structure of the plaid model can look similar for different parameters.

We consider the two rationals $p_1/q_1 = 5/12$ and $p_2/q_2 = 12/29$. These rationals are related in a Diophantine sense. They are consecutive continued fraction approximations to $\sqrt{2} - 1$. Moreover, we note that

$$\alpha_1 = \tau_1/\omega_1 = 5/17, \quad \alpha_2 = \tau_2/\omega_2 = 17/41. \quad (17)$$

We have $\alpha_2 \approx \alpha_1$. For this reason, the lines of low capacity for the two parameters are almost exactly in the same relative positions, as are the lines of low mass. Hence, the low mass light points are distributed about the same way for both parameters. Were we to use the low capacity information for the parameter $12/29$ as we did for the parameter $5/12$, we would see almost exactly the same picture. This explains the large scale resemblance between the two pictures. Of course, when we start looking at high capacity lines we get down to the fine details of the pictures and they look very different.

Figure 3.5 is just an example of a fairly ubiquitous phenomenon. We will discuss the general case below, after we make the grid method above more robust.

3.3 Subdivision Algorithm

The approach above gives us a way to see some of the coarse structure of the plaid model. However, one shortcoming of the method is that there might not be any (or many) good triples associated to a parameter. As we increase K , the number of rectangles increases, and one ambiguous rectangle ruins the grid. However, we can remedy the problem by subdividing an ambiguous rectangle into smaller rectangles using lines of higher capacity. With this approach, we keep the good part of the coarse grid and refine only as needed. We will illustrate this with an example which occurs in the triple $(38/161, B, 3)$ for some block B .

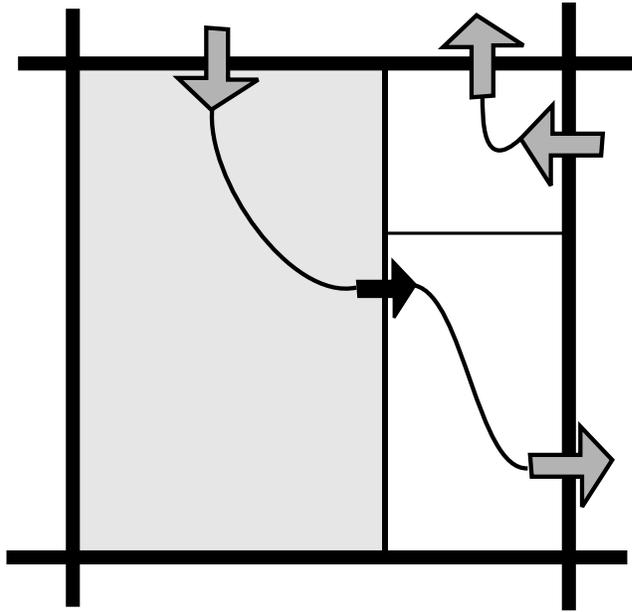


Figure 3.6: Resolving an ambiguous rectangle

The outer rectangle in Figure 3.6 belongs to Γ_3 , and the smaller rectangles belong to Γ_K for a somewhat larger value of K . We find the lowest capacity vertical line which divides crosses the big rectangle. This line has only one light point inside the big rectangle. Now we have an unambiguous rectangle (shaded) and a tall thin ambiguous one. Next, we put in the lowest capacity horizontal line which crosses the ambiguous small rectangle. There are no light points on this segment inside the relevant rectangle, and so the ambiguous rectangle resolves into two unambiguous ones.

Typically this subdivision algorithm terminates quickly. However, if we carefully choose the blocks and the parameters, we can make it go as long as we like. Figure 3.7 shows a more complicated example in which it takes much longer to resolve an ambiguous rectangle. We have not completed the resolution here. The two shaded rectangles are still ambiguous. We have somewhat indicated the order in which the subdivision takes place by the thickness of the lines. The thicker lines are added first.

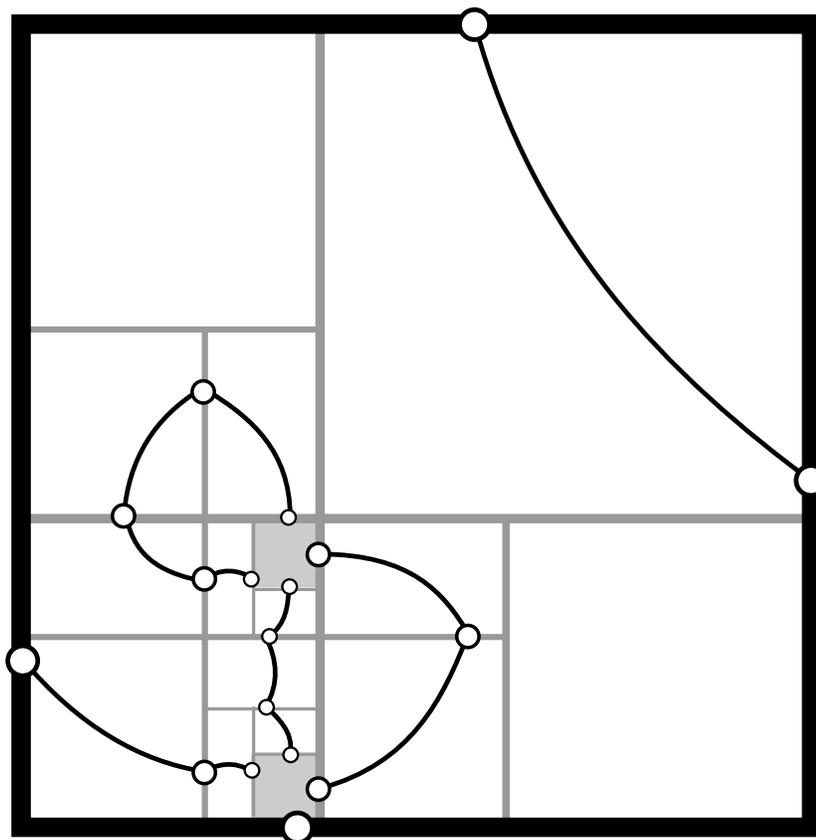


Figure 3.7: Resolving an ambiguous rectangle

Since the plaid model is consistent, the subdivision algorithm always terminates in a completely unambiguous pattern of rectangles. However, if we end up with the final grid of 1 by 1 squares we haven't really obtained useful coarse information about the plaid model.

3.4 Grid Lines as Barriers

In this section we will use the model to get some information about the distribution of the orbits. This material will be useful for Theorem 0.7.

Fixing a parameter p/q and a block B and an even integer $K \geq 0$ let Γ_K denote the union of all the lines of capacity at most K which intersect B . The complement $B - \Gamma_K$ consists of $(K + 1)^2$ rectangles arranged in a grid pattern. We say that one of these rectangles is *empty* if its boundary has no light points on it. Empty rectangles serve as barriers, separating the plaid polygons inside them from the plaid polygons outside them.

Lemma 3.2 (Empty Rectangle) *For all parameters, all blocks B , and all choices of K , at least one of the rectangles of $B - \Gamma_K$ is empty.*

Proof: We suppose there are no empty rectangles and derive a contradiction. There are a total of $(K + 1)^2$ rectangles. If some rectangle R has a light point on it, then it must have a second light point, because the polygon Γ crossing into R through an edge containing one of the light points must cross out of R through another edge.

The one exceptional situation is when the light point z lies at the corner of R . In this case, one of the edges E of R lies in a vertical boundary of the block B . Let's consider the case when E lies in the west boundary of R and z is the south west corner. The other cases are similar. Γ crosses into R through the south edge of R , but then it cannot exit through E because E lies in the boundary of B . Even in this exceptional case, there must be 2 light points in the boundary of R .

If every rectangle has at least 2 light points, then there are at least $(K + 1)^2$ light points total. On the other hand, we know that a line of capacity k contains at most k light points. Since there are 4 lines of capacity k for each $k = 0, 2, \dots, K$, this gives a total of

$$8 \sum_{k=1}^{K/2} k = (K + 1)^2 - 1.$$

We have one fewer point than we need. This is a contradiction. ♠

In practice, we see many empty rectangles within a block. One consequence of the Empty Rectangle Lemma is that a plaid polygon cannot be

nearly dense in a block. It is excluded from at least one rectangle. When the Empty Rectangle is used in a recursive way, for many grids, it guarantees the existence of many distinct plaid polygons within a block. We will elaborate on this when we prove Theorem 0.7.

We can use the count from the Empty Rectangle Lemma to give us some information about the sizes of the polygons in the plaid model. Recall that $\omega = p + q$. For any plaid polygon P , define

$$\delta(P) = \frac{\max(\|P\|_1, \|P\|_2)}{\omega}. \quad (18)$$

Here $\|P\|_j$ is the diameter of the projection of P onto the j th coordinate direction. We always have $\delta(P) < 1$. The big Polygon Γ has the property that $\delta(\Gamma) > 1/2$.

Given any K , let $\epsilon'(K)$ denote the maximum spacing between a pair of adjacent lines in the grid Γ_K , and let $\epsilon(K) = \epsilon'(K)/\omega$. Were we to rescale the blocks to have unit size, $\epsilon(K)$ would denote the largest dimension of one of the rectangles cut out by Γ_K .

Theorem 3.3 *In any block and for any K , there are less than $(K + 1)^2/2$ plaid polygons P for which $\delta(P) > \epsilon(K)$.*

Proof: If $\delta(P) > \epsilon(K)$ then P must cross some line of Γ_K . But then P must contain at least 2 of the light points associated to Γ_K . We have already seen that there are less than $(K + 1)^2$ such grid points. ♠

Example: Lemma 3.3 is most useful when we have some control on the geometry of the grid Γ_K . If the lines of Γ_K are well equidistributed, then $\epsilon(K) = O(1/K)$. This depends on the properties of τ/ω . The good case is when the terms in the continued fraction expansion of τ/ω are small. For instance, consider the sequence of numbers $\{a_n\}$ where

$$a_1 = 1, \quad a_2 = 2, \quad a_{n+1} = 2a_n + a_{n-1}. \quad (19)$$

The sequence starts out

$$1, 2, 5, 12, 29, 70, \dots$$

Define $p_n = a_{2n+1}$ and $q_n = a_{2n+2}$. The sequence of rationals starts out

$$1/2, 5/12, 29/70, \dots,$$

and $p_n/q_n \rightarrow \sqrt{2} - 1$. One can check inductively that

$$\omega_{n-1}\omega_n = 2p_n^2 + 1.$$

Hence $\tau_n = p_{2n+1}$ and $\tau_n/\omega_n \rightarrow \alpha = 1 - 1/\sqrt{2}$. The continued fraction expansion of α is

$$0 : 3 : 2 : 2 : 2 : \dots$$

With some effort, one can check that the largest gap in the spacing of the lines of Γ_K is less than 3 times the smallest gap. This gives $\epsilon(K) < 3/K$. This bound holds relative to any p_n/q_n as long as $K < \omega_n/2$. In all these cases there are less than $(K + 1)^2/2$ polygons P with $\delta(P) > 3/K$.

4 Three Dimensional Construction

4.1 Remote Adjacency

The results in the Section 2.5 are a kind of conservation principle. As we move from block to block, the number of light points on a given line does not change. In this section we explain how to think about these points as moving particles. This point of view turns out to be crucial to the three dimensional interpretation of the plaid model.

Define

$$B_k^\pm = [k\omega, k\omega + \omega] \times [0, \pm\omega]. \quad (20)$$

Here $B_0^+, \dots, B_{\omega-1}^+$ are the fundamental blocks. In the undirected plaid model, the two blocks B_k^+ and B_k^- have exactly the same picture in them. In the directed model, these blocks have the same picture except that all the orientations have been reversed. We write

$$B_k^\pm \rightarrow B_{k-\hat{\tau}}^\mp \quad (21)$$

We call the two blocks involved in this relation *remotely adjacent*. Even though these blocks are far apart in the plane, we will make the case in §4 that they should really be stacked on top of each other in a 3 dimensional model.

Since $\hat{\tau}$ is relatively prime to ω , the cycle

$$B_0^+ \rightarrow B_{-\hat{\tau}}^- \rightarrow B_{-2\hat{\tau}}^+ \rightarrow B_{-3\hat{\tau}}^- \rightarrow \dots \quad (22)$$

encounters every one of the blocks defined above. We call this the *fundamental cycle*. It has length 2ω . As we go through the fundamental cycle, we imagine referring points in the various blocks back to the first block by translation. What we are interested is the relative position of a point within the block. Given $z \in B_k^\pm$ we define $[z] \in B_0^+$ to be the image of z under the translation which carries B_k^\pm to B_0^+ .

Remark: This point of view is very natural on our computer program. One can move through the fundamental cycle and the program will automatically recenter the drawing window on the current block. Thus, what we naturally see in the program as we go through the fundamental cycle is the *relative motion* of the light points.

4.2 Horizontal Particles

Lemma 4.1 *Suppose $B \rightarrow B'$. Let H be a \mathcal{H} line which intersects B and B' . Let z be an intersection point on H . Suppose z has type P and is not on the right edge of B . Then there is a point $z' \in B' \cap H$ of the same type and shade as z such that $[z'] - [z] = (P^{-1}, 0)$. If z is a light point, then the direction of z' is also the same as the direction of z .*

Proof: Let L be the slanting line of slope $-P$ that contains z . Define

$$z' = z + (-\hat{\tau} + P^{-1}, \pm\omega). \quad (23)$$

The sign in the second coordinate is chosen according to the parity of B . Since z is not on the right edge of B and the slanting line through z has slope $-P$, we see that z is at least P^{-1} from the right edge of B . Hence $z' \in B'$. By construction, z' still lies on a slanting line L' of slope $-P$. The difference between the y -intercepts of L and L' is $-2P\tau + 1 + \omega$. This number is even and congruent to $0 \pmod{\omega}$. Hence it is also $0 \pmod{2\omega}$. Hence L and L' have the same signed mass and direction. In case z is a light point, the directions of z and z' are determined by the directions of L and L' . From the argument in Lemma 1.3 these lines have the same direction. Hence, z and z' have the same direction. ♠

The same argument proves the following result.

Lemma 4.2 *Suppose $B \rightarrow B'$. Let H be a \mathcal{H} line which intersects B and B' . Let z be an intersection point on H . Suppose z has type Q and is not on the left edge of B . Then there is a point $z' \in B' \cap H$ of the same type and shade such that $[z'] - [z] = (-Q^{-1}, 0)$. If z is a light point, then the direction of z' is also the same as the direction of z .*

If z and z' are related as in the two lemmas above, we write $z \rightarrow z'$. We call the collection of points in the cycle

$$z \rightarrow z' \rightarrow z'' \rightarrow \dots \rightarrow z^{(2\omega-1)} \quad (24)$$

a *horizontal particle*. We call each of the points in the cycle an instance of the particle. If we go through the fundamental cycle twice and watch the *relative positions* of the instances of a particle, the points appear to move rightward with speed P^{-1} and leftward with speed Q^{-1} . The reader can see this in action using my computer screen.

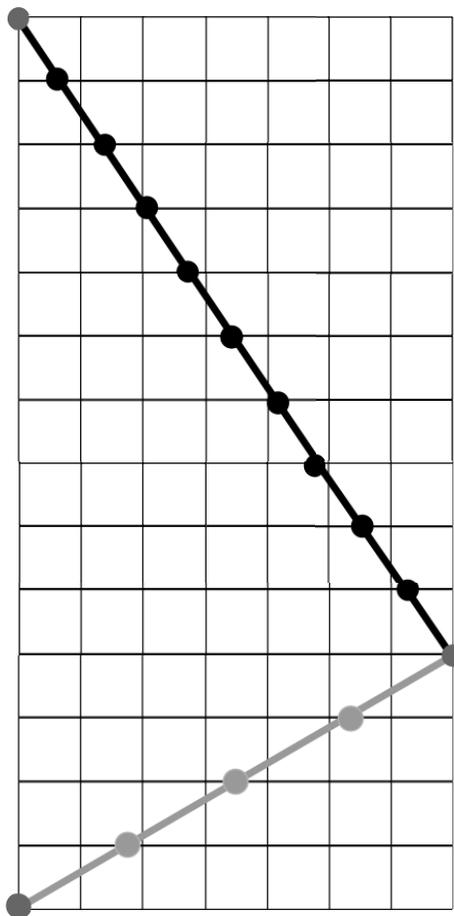


Figure 4.1: A spacetime diagram for a horizontal particle

We can draw a *spacetime diagram* showing the successive points of the particle. The bottom row of the diagram shows the location of z on the horizontal segment that contains it. The second row shows the location of z' on the segment that contains it, and so on. In the example shown in Figure 4.1, the parameter is $2/5$ and the horizontal line has capacity 2. We are tracking the particle whose 0th instance is the light point $(0, 2)$ of mass 1. The top and bottom point are meant to be identified, so that there are $14 = 4 + 10 = 2 \times 7$ instances in total. The slopes of the lines are $P = 4/7$ and $-Q = -10/7$. All other diagrams for the horizontal light points for the same parameter look the same up to translation.

We will have quite a bit more to say about these spacetime diagrams below.

4.3 Vertical Particles

In the vertical case, the situation is superficially different because each vertical line intersects at most one fundamental block. So, we consider a family of vertical lines $\{V_k\}$ such that V_k intersects the k th block in the fundamental cycle in the same relative position. The following result is proved the same way as in the horizontal case.

Lemma 4.3 *Let $B \rightarrow B'$ be fundamental blocks. Let $V \rightarrow V'$ be vertical lines which respectively intersect these blocks. Let z be an intersection point of type P (respectively type Q) on V . Then there is an intersection point $z^* \in V'$ of the same type and shade so that $[z^*] = [z] + (0, 1)$ (respectively $[z^*] = [z] - (0, 1)$.) If z is a light point then the directions of z and z^* are the same.*

There is one subtle point about Lemma 4.3. If the point z has type P and is in the top unit vertical segment on V , then the point z^* does not actually lie in the block Z^* . We can attempt to fix this problem by taking the point $z^* - (0, \omega)$ instead. However, when z^* is a light point, the direction of $z^* - (0, \omega)$ is opposite that of z . This is something we don't like. Also, the points z and $z^* - (0, \omega)$ are not very close to each other relatively speaking.

There is a saving grace in this case. Let ρ denote the horizontal reflection in the horizontal midline of the block B' . In case z has type P and lies in the top unit integer segment of B we define

$$z \rightarrow z' = \rho(z^* - (0, \omega)). \quad (25)$$

By symmetry, z' is a light point having the same direction as z . The two points $[z]$ and $[z']$ both lie in the top unit integer segment of the relevant grid line, and the sum of their distance to the top of the block is 1 unit. It is as if $[z]$ has bounced off the top wall and arrived at $[z']$. The spacetime diagram below will make this more clear.

When z has type Q and lies in the bottom unit integer segment, we write

$$z \rightarrow z' = \rho(z^* + (0, \omega)). \quad (26)$$

The same remarks in this apply.

In all cases, we define $z^* = z'$. Thus we get a cycle just as in Equation 24. When the particle is, in a relative sense, traveling up, it has type P. When it is traveling down, in a relative sense, it has type Q. We call the union of 2ω light points in the cycle a *vertical particle*.

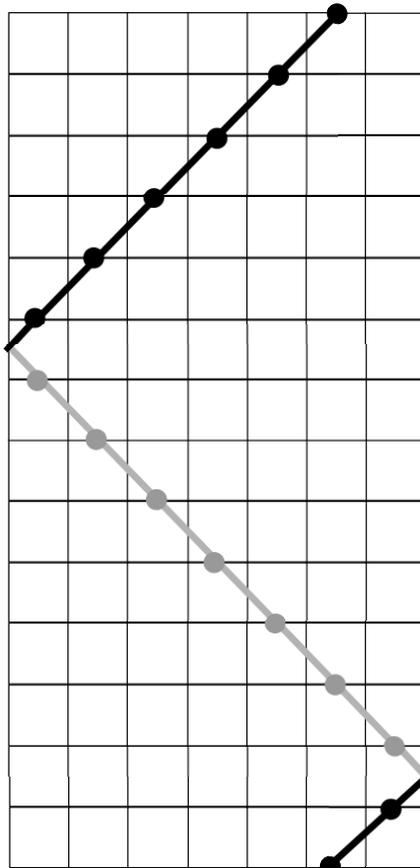


Figure 4.2: A spacetime diagram for a vertical particle.

We can make a spacetime diagram as in the horizontal case. This time we turn the vertical grid lines horizontal and stack them on top of each other. Figure 4.2 shows the diagram for one of the light points on a vertical line of capacity 6 for the parameter $2/5$. As in the horizontal figure, there are $14 = 7+7$ instances of the particle in this example. The lines in this case have slope 1. Remembering that the top and bottom of the spacetime diagram are identified, we see that what we have is a billiard path in a cylinder. The instances along this path change type when the path bounces off the side walls of the cylinder.

We chose to show the diagram for this particle because the instances of the particle lie fairly near the centers of the unit integer segments that contain them. This clearly illustrates our point that, in the vertical case, particles never actually lie on the top or bottom edge of the block.

4.4 Stacking the Blocks

Now we take the idea of spacetime diagrams farther. We work with the directed model, as above. We fix a parameter p/q throughout the discussion and set $\omega = p + q$ as usual. The 2ω fundamental blocks all lie in the plane, but we can cut them out of the plane (so to speak) and stack them so that remotely adjacent blocks, in the sense of §4.1 are consecutive integer slices of the rectangular solid

$$\Omega = [0, \omega]^2 \times [0, 2\omega]. \quad (27)$$

To be more precise, the block $B_{-k\hat{\tau}}^\pm$ is identified with the horizontal slice $[0, \omega]^2 \times \{k\}$ for $k = 0, \dots, 2\omega - 1$. In other words, we are using the extra dimension to make remotely adjacent slices actually adjacent.

To make the picture nicer, we think of Ω as a Euclidean orbifold, with the following markings:

- The face $z = 0$ is identified by translation to the face $z = \omega$.
- The remaining faces are mirrored.

These identifications guarantee that the bottom slice and the top slice are also spaced 1 apart in the orbifold. They also capture the “bounce” properties of the spacetime diagrams.

The horizontal spacetime diagrams described above are really just XZ slices of Ω . Likewise the vertical spacetime diagrams are really just YZ slices of Ω . Technically speaking, the lines in these diagrams are just guides for the eye to follow the particle. However, in the sections following this one, we will enhance our view of the plaid model so that the lines, in a sense, become part of it.

In the spacetime diagrams above, we just showed one particle at a time. However, we show all the particles associated to a given grid line at the same time. If the grid line has capacity k , there will be k translated copies of the polygonal paths shown in Figures 4.1 and 4.2. Figure 4.3 shows the diagrams for a horizontal slice of capacity 2 and or a vertical slice of capacity 2. We have used the parameter $2/5$ again.

The spacetime slices are really cylinders in the orbifold Ω . The tops and bottoms are identified. For this reason, the diagrams in Figure 4.3 actually have more symmetry than first meets the eye. If the cylinder is cut open and redrawn with a different “bottom”, then the picture looks nicer. Figure 4.4 shows the nicer drawing.

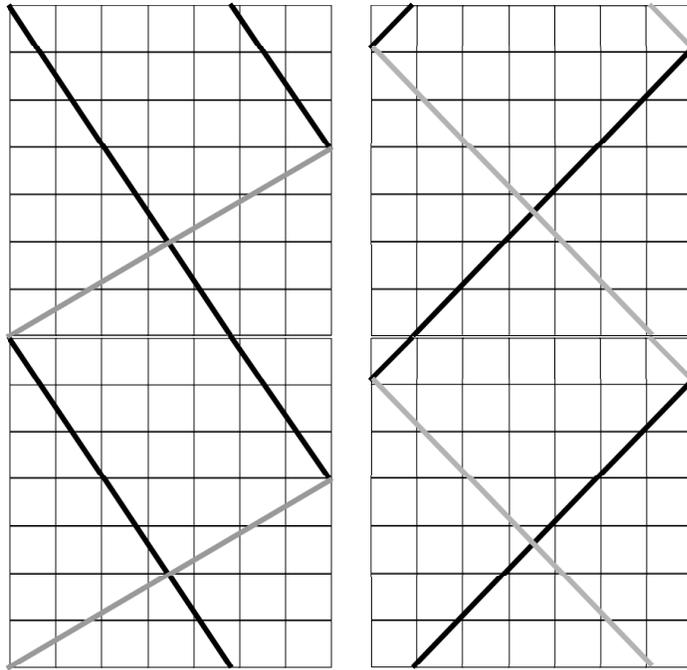


Figure 4.3: Capacity 2 spacetime diagrams for the parameter $2/5$.

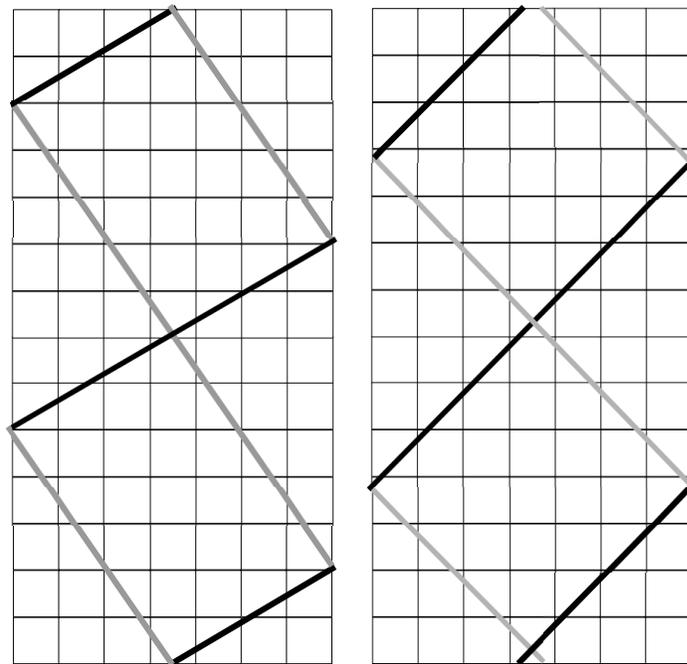


Figure 4.4: Improved capacity 2 spacetime diagrams for the parameter $2/5$.

4.5 Pixellated Spacetime Diagrams

Now we take the next step in defining a 3 dimensional enhancement of the plaid model. We can convert each planar spacetime diagrams into a union of tiles, where each tile has either 0, 1, or 2 non-crossing segments joining the centers of the edges of the tile. In each square we connect the centers of the edges which are crossed by an odd number of particle lines. Since each particle line enters and exits a square, there is automatically an even number of such odd crossings. In case there are 4 odd crossings, we pair those which come from the same particle line. We call this process *pixellation*. Figure 4.5 shows the pixellation of Figure 4.4. The pixellation rules allow for the particles to merge and lose their individual identities. In Figure 4.5 each pixellated path is a closed and centrally symmetric loop in the orbifold.

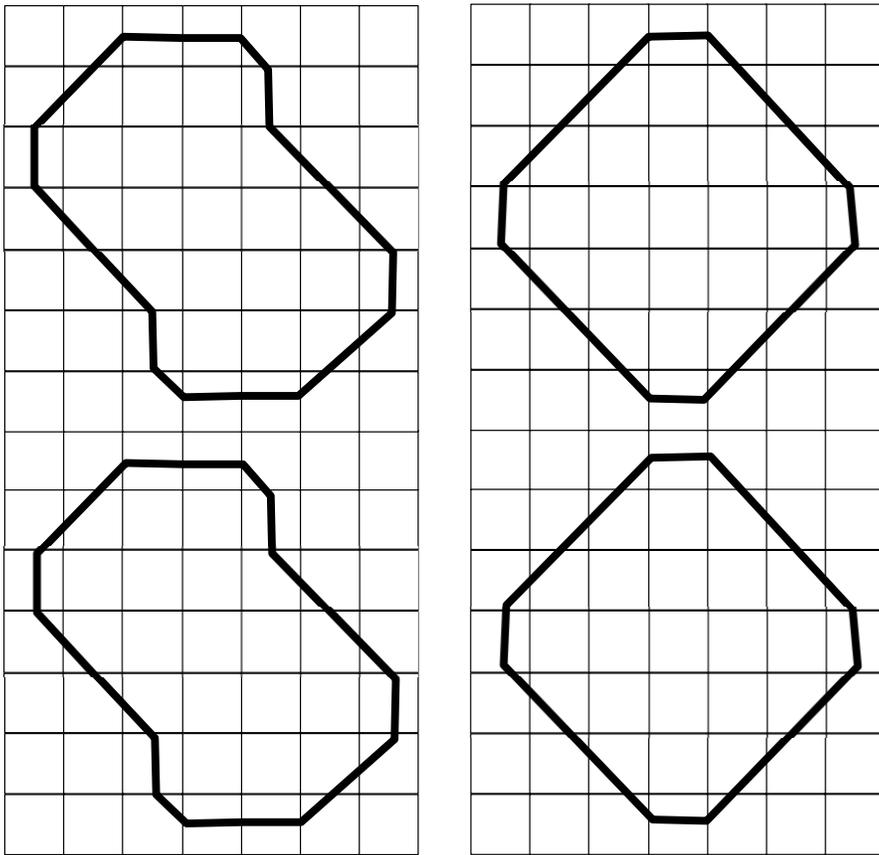


Figure 4.5: pixellated spacetime diagrams

Figure 4.6 shows a more complicated example, taken with respect to the parameter $7/16$. The left side shows the spacetime diagram for one of the capacity 6 horizontal lines and the right hand side shows the spacetime diagram for one of the capacity 6 vertical lines.

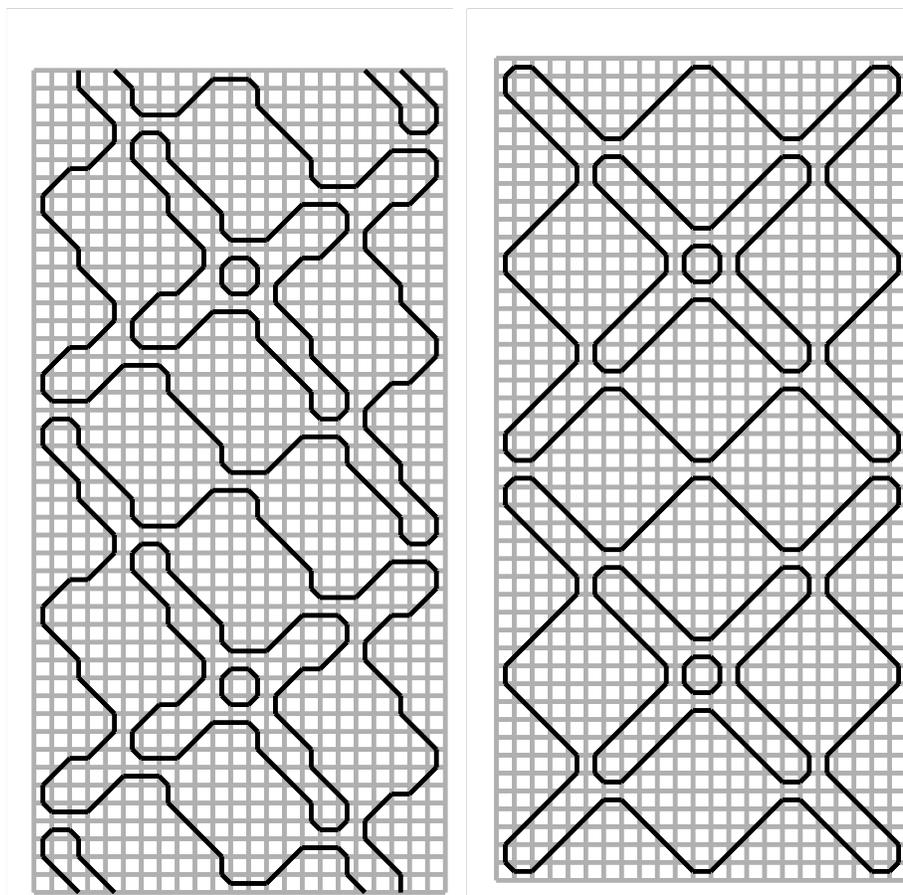


Figure 4.3: Pixellated spacetime diagrams for $\kappa = 6$ and $p/q = 7/16$.

Notice that the two curve families are isotopic. Relative to the parameter p/q , the horizontal and vertical pixellated spacetime diagrams associated to the capacity κ are always isotopic when $\kappa \leq 2p$. This fact is a byproduct of our proof of the Truchet Correspondence Theorem.

4.6 Spacetime Plaid Surfaces

By construction, a tile in an XY plane is compatible with a tile in the XZ plane, in the following sense. If two tiles share an edge, then either both connectors avoid this edge, or both connectors use the edge. This is to say that the two tiles fit together in a compatible way. The idea is just that the common edge is a horizontal edge of a unit integer square, and this edge either has a light point or not. The same compatibility holds for tiles in the XY plane and in the YZ plane which are adjacent.

Now comes the interesting part. It turns out that the tiles in the XZ plane are compatible with the tiles in the YZ plane in the same sense. We will prove this momentarily, but first we explain the main consequence. Consider a unit integer cube in the spacetime picture. On each face of the cube we have a tile. these tiles are completely compatible with each other across edges. this means that the union of the connectors on the tiles is a union of closed embedded polygons on the surface of the cube!

Each such polygon bounds a disk inside the cube. If we like, we can take these disks to be polyhedral disks, though generally they are not planar. Call the union of these polyhedral disks a *filling* of the cube. Figure 4.7 shows some examples which actually arise. The thick black edges are parts of plaid polygons and the thick gray edges are parts of particle lines.

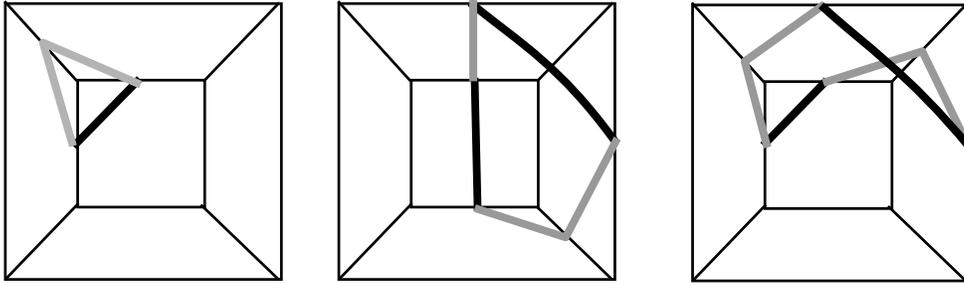


Figure 4.7: Some typical cubical fillings

Our computer program lets the user see the cubes and their fillings. If we put in fillings for every polyhedral cube, then the result is a finite union of pairwise disjoint polyhedral surfaces. If we slice the surfaces in any integer XY plane, we get the plaid polygons in the corresponding block. If we slice the surfaces in any XZ or YZ integer plane, we get the corresponding pixellated spacetime diagrams. Thus the pixellated spacetime diagrams and the plaid polygons are all just slices of what we call *plaid surfaces* in the

spacetime interpretation of the plaid model. We don't have any specific uses for these surfaces, but we think that it shows off the depth and beauty of the plaid model.

Theorem 4.4 *The tiles in the XZ plane are comparable with the tiles in the YZ plane. That is, if two such tiles share a common edge, then this edge is involved in the one connector if and only if it is involved in the other.*

Proof: Suppose that a given edge E is involved in a connector of an XZ tile τ . There is some $t \in \{0, \dots, \omega - 1\}$ such that E connects a horizontal edge at time t to a horizontal edge at time $t + 1$ in the relevant XZ plane Π . In the planar picture, a single horizontal light point in some horizontal line L_1 crosses over a vertical line L_2 as we move from B_t to B_{t+1} . The intersection point $L_1 \cap L_2$ corresponds to the edge E .

Let's compare the union of 4 squares $B_t - (L_1 \cup L_2)$ and $B_{t+1} - (L'_1 \cup L'_2)$. Here L'_j is the translate of L_j by $(-\omega\hat{\tau}, 0)$ so that it lies in B_{t+1} . Because the plaid model is coherent, there must be an even number of light points in the boundary of each of the 8 squares mentioned above. Hence, as our particle v crosses over L_2 there must be a second particle v' which crosses one of the two lines L_1 or L_2 . There are two cases to consider.

If v' crosses L_2 then the corresponding particle line also crosses E in τ . Given the slopes and spacing of the particle lines, there is at most 2 particle lines crossing τ . Hence, there are exactly 2. But then E is not involved in the connector of τ . This is a contradiction.

If v' crosses L_1 then v' is a vertical particle. In this case, the corresponding particle line in the relevant YZ plane crosses E . If a second line in the YZ plane crossed E then a second vertical particle v'' would also be crossing over L_2 from the other direction. But then the parity would be off in the discussion above. There would have to be a second horizontal particle crossing over L_2 . So, there is no v'' and we see that the connector in the tile τ' involves the edge E . ♠

Theorem 4.4 guarantees the existence of the surfaces in the spacetime version of the plaid model. Here is some of the combinatorial structure. Each surface is closed, and tiled by polygons, all of which have at most 8 sides. The polygons fit together 4 around every vertex. The 1-skeleton of the tiling is colored according to the kind of slice containing the edge.

4.7 The Simplest Example

Here we explain the structure of the plaid surfaces in the simplest case, when $p/q = 1/2$. It is convenient to work with the undirected model, which is just the 2-fold quotient of the directed model. In the undirected version, there are 3 fundamental blocks, and the plaid polygons are shown in Figure 4.8.

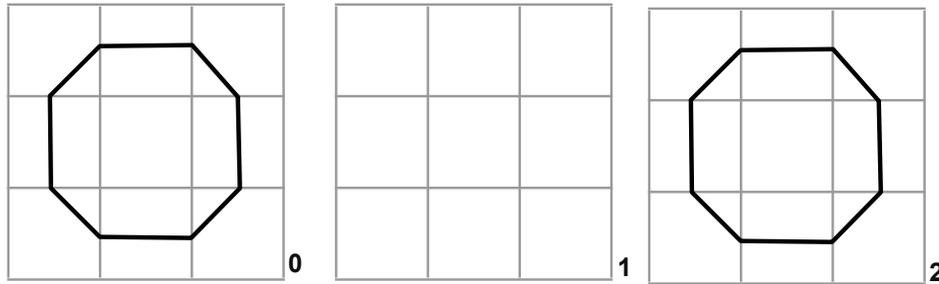


Figure 4.8: The three blocks for the parameter $1/2$.

There is exactly one surface in Ω . The surface is a topological sphere that is perhaps best described by starting with a Rubik's cube, shaving off half of each edge, and taking the boundary. Figure 4.9 shows a topologically accurate planar picture. The thick lines make the plaid polygons. The thin lines come from the pixellated worldlines. If we forget the coloring of the lines, the surface has a symmetry group of order 48.

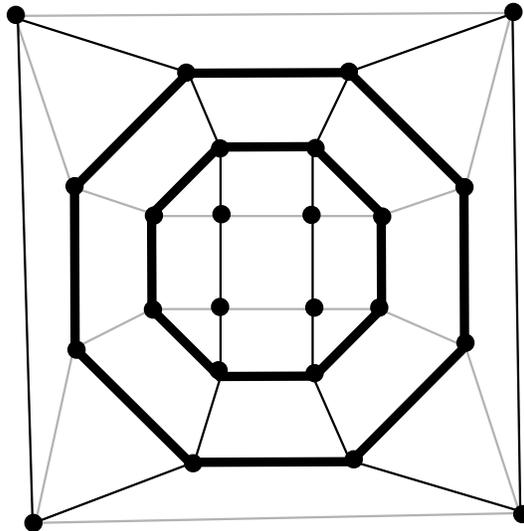


Figure 4.9: The spacetime surface for the parameter $1/2$.

4.8 A Remark about Rescaling Limits

The spacetime picture gives us a natural framework in which to take rescaled limits. We rescale the flat tori so that they all have unit diameter. Suppose we have two parameters p_1/q_1 and p_2/q_2 where $\min(q_1, q_2)$ is large. Suppose also that the auxiliary ratios α_1 and α_2 defined in Equation 17 either satisfy $\alpha_1 \approx \alpha_2$ or $\alpha_1 \approx 1 - \alpha_2$. Here $\alpha_j = \widehat{\tau}_j/\omega_j$. If we fix a capacity cutoff κ that is much smaller than ω , we would expect that the subdivision algorithm would terminate quickly if we place a low capacity limit on the grids we consider, yielding combinatorially identical approximations to the plaid polygons in each XY slice.

At the same time, we would expect the low capacity pixellated spacetime diagrams to be combinatorially identical. (What is missing here is an understanding of how the auxiliary orientations are produced.) All in all, the plaid surfaces ought to look about the same in a number of slices in all 3 coordinate directions,

All this geometry ought to make it possible to take a rescaled limit given a sequence $\{p_n/q_n\}$ of rational parameters in which the auxiliary quantities $\{\alpha_n\}$ also converge. The limit ought to be a flat torus filled with fractal surfaces at all scales. I have yet to work this out.

5 Directed Spacetime Diagrams

In this chapter we will prove a technical result, the Curve Turning Theorem, which helps us understand the structure of the spacetime diagrams. In the next chapter we will use the Curve Turning Theorem to prove the Truchet Correspondence Theorem, the result which connects the plaid model to Truchet tilings.

5.1 The Basic Definition

So far, we have neglected a key feature of the spacetime diagrams. They are directed. We took the trouble to prove that the directions of all instances of a particle are the same, and now we are going to use this fact.

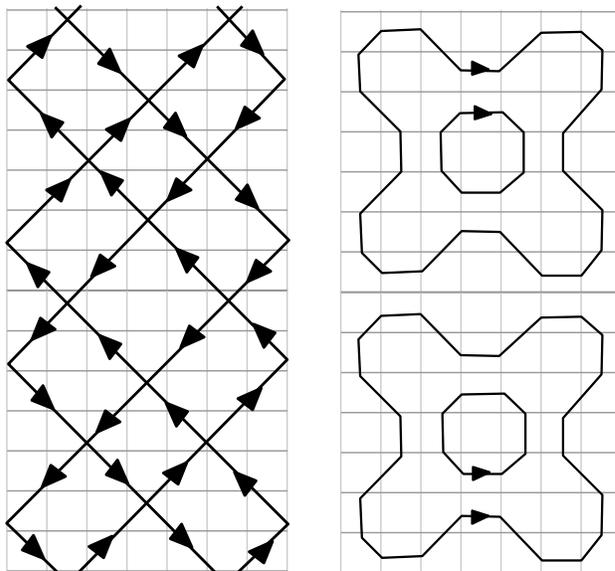


Figure 5.1: A directed vertical spacetime diagram and its pixellation

We define a *worldline* to be a polygonal path which describes a particle. We direct a horizontal worldline upward if the transverse direction points upward. We direct a vertical worldline upward if its transverse direction points to the right. The choice of up/down versus left/right is not important. Were we to pick a different convention, all the directions within a diagram would be reversed and we would get the same overall result. The left side of Figure 5.1 shows an example of a vertical capacity 4 spacetime diagram for the parameter $2/5$.

5.2 The Curve Turning Process

The right side of Figure 5.1 shows the pixellation of the left side. In this case, the pixellation process has an alternate description: We follow along the directed segments of the diagram and turn at every intersection. We call this the *curve following process*. Technically, we do not trace directly on top of the curves, but rather move slightly so that the paths we draw connect up the centers of the unit integer squares. In Figure 5.1 the curve turning process and the pixellation process produce the same result. This is no coincidence.

Theorem 5.1 (Curve Turning) *For any even rational parameter p/q and any capacity $\kappa \in \{2, 4, \dots, 2p\}$, let Σ be any of the corresponding directed spacetime diagrams of capacity κ . Then the pixellation process and the curve turning process produce the same set of polygons.*

One reason we are interested in this result is that we will use it to establish the connection between the plaid model and the Truchet tile system. The second reason is that the Curve Turning Theorem points out the hidden robustness of our spacetime surfaces. The pixellation process looks like it is very delicate. If we jiggle the lines a bit, perhaps we get a completely different pattern. On the other hand, the curve-following process is quite robust. It only depends on the topological pattern in which the directed lines intersect. (Compare Figure 6.3 in the next chapter.) Here is a sample corollary.

Corollary 5.2 *The following is true for any even rational parameter. There is only one plaid surface in the undirected spacetime model which intersects the capacity 2 slices, and this surface intersects each slice in a single polygon.*

Proof: The corresponding directed spacetime diagram contains two directed polygonal paths. Within each block, the corresponding light points correspond to a single directed plaid polygon and hence have opposite directions. Thus, we get the same picture as on the right hand side of Figure 5.1. By the Curve Turning Theorem, the corresponding pixellated spacetime diagram always has 2 translation-equivalent loops. When we consider the undirected model, these loops are identified. The polygon from §3.1 intersects all 4 capacity 2 slices and so the same surface intersects all the capacity 2 slices. ♠

Remarks:

(i) Corollary 5.2 allows us to say that there is a fundamental surface, for each parameter, in the undirected spacetime diagram. It is the surface which intersects the lowest capacity slices. This points out some additional significance of the fundamental polygon from §3.1. It is the bottom slice of the fundamental surface.

(ii) A somewhat more sophisticated analysis would show that the pattern of intersection for the capacity 4 slices is always as in Figure 5.1 for any parameter p/q with $p \geq 2$. Thus, the union of plaid surfaces in the undirected model intersects the capacity 4 slices in two nested polygons. The outer nested polygon is always a slice of the fundamental surface. See Theorem 6.2.

5.3 Two Exclusion Principles

The material in this section is preliminary to the proof of the Curve Turning Theorem. We work with the directed plaid model.

Lemma 5.3 (Exclusion I) *A unit integer line segment cannot contain two light points having opposite types and the same directions.*

Proof: Call the points ζ_1 and ζ_2 . Consider the vertical case. Suppose ζ_1 lies on a slanting line of slope $-P$ and ζ_2 lies on a slanting line of slope $-Q$. Define the *separation* between two light points on the same vertical line to be the distance between the centers of the vertical unit integer segments containing them. Let ζ'_1 be the reflection of ζ_1 in the horizontal midline of the relevant block. By symmetry, ζ'_1 and ζ_1 have opposite directions. The separation between ζ'_1 and ζ_1 is even, by symmetry. So, it suffices to prove the separation between ζ'_1 and ζ_2 is odd.

ζ'_1 and ζ_2 have opposite signs and both lie on slanting lines of the same negative slope. Hence, their distance is the same as the distance between the y -intercepts of these two slanting lines. By the argument in Lemma 1.3, the distance between these two y -intercepts is an odd integer. Hence, the separation between ζ'_1 and ζ_2 is odd.

In the horizontal case, ζ_1 lies on slanting lines of slope $\pm P$ and ζ_2 lines of slope $\pm Q$. Let $(m + 1/2, y)$ denote the midpoint of the horizontal unit segment containing ζ_1 and ζ_2 . By symmetry it suffices to consider the case

when $m \geq 0$. Let d_1 and d_2 be the distance from ζ_1 and ζ_2 respectively to this midpoint. Note that $d_1, d_2 \in (0, 1/2)$.

Let y_1 denote the y -intercept of the slanting line L_1 of slope $-P$ which contains ζ_1 . Let y_2 denote the y -intercept of the slanting line L_2 of slope Q which contains ζ_2 . Since $P + Q = 2$, we have

$$|y_1 - y_2| = (P + Q)(m + 1/2) \pm Pd_1 \pm Qd_2 = 2m + 1. \quad (28)$$

The two slanting lines of the same slope through ζ_1 and ζ_2 have the same sign. So, from the second definition of the plaid model, L_1 and L_2 have opposite sign. So, by the argument in Lemma 1.3, L_1 and L_2 are both directed the same way. Since these lines have opposite slope, one of them points down at the relevant light point and the other one points up. Hence ζ_1 and ζ_2 have the opposite direction, a contradiction. ♠

Remark: The same argument shows that we cannot have two vertical light points having opposite type, opposite direction, and odd separation.

Now we prove a second exclusion principle. the even parameter p/q is implicit in the result.

Lemma 5.4 (Exclusion II) *It is impossible to have two consecutive slanting lines of the same slope and opposite direction, both of which have mass less than $2p$.*

Proof: Let y and $y + 1$ have be the y -intercepts of these two lines. Since y and $y + 1$ have opposite parity, the slanting lines through these points have the same sign. The capacities are given by the representatives of $2py$ and $2py + p \bmod 2\omega$ which lie in $(-\omega, \omega)$. The difference between these two representatives is congruent to $2p \bmod 2\omega$. But it is impossible to have two integers of the same sign in $(-2p, 2p)$ whose difference is congruent to $2p \bmod 2\omega$. ♠

5.4 Proof of the Curve Turning Theorem

Before we launch into the proof, we remark that the horizontal case is much more painful to prove, even though the ideas are the same. I discovered the

tedious proof in the horizontal case after a hard struggle. The vertical case is sufficient to establish a connection between the plaid model and the Truchet tile system, and one could view the horizontal case as icing on the cake. For better or worse, here is the proof of both cases.

It never happens that there are two directed segments in Σ which are within 1 unit of each other, parallel, and pointing in opposite directions. Such segments would correspond to particles having some instances which lie on consecutive slanting lines of the same negative slope and opposite directions, and this would contradict the Exclusion Lemma II. The existence of such lines is the only way it can happen that the pixellation process gives tiles which have more than one connector in them. So, the pixellated tiles have at most 1 connector in them.

Since there is only one connector in each tile, the situation is easy to analyze in a unit square which only has one worldline going through it. In this case, the intersection of the worldline can be moved by at most $\sqrt{2}/2$ pointwise so that it matches the pixellated version.

We just have to worry that the two processes do opposite things at an intersection point. First consider the case when two up-pointing worldlines intersect, as shown in Figure 5.2.

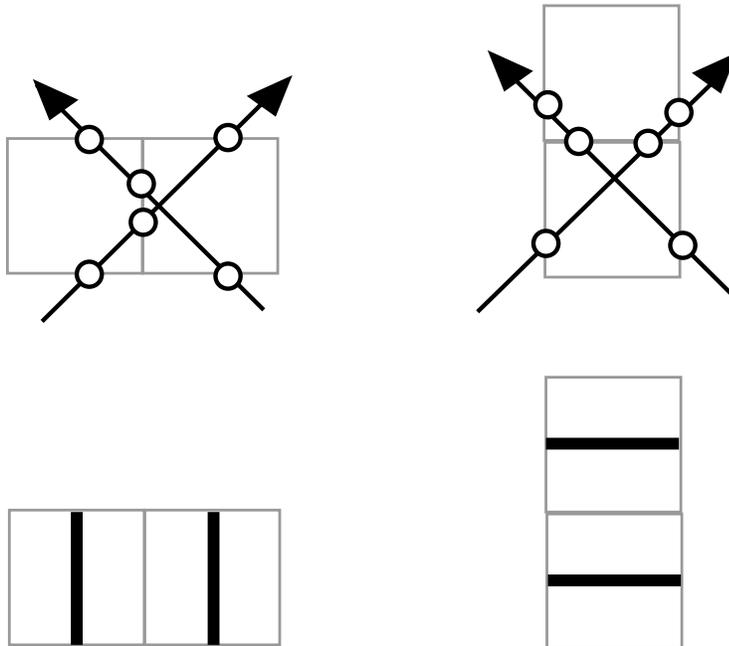


Figure 5.2: two worldlines intersecting

Of the two possibilities shown in Figure 5.2, the one on the right is ruled out by the Exclusion Lemma 1. So, for intersections between up-pointing worldlines, the two processes yield the same results. The same goes for the intersection of two down-pointing worldlines.

Now consider the case when the worldlines intersect and are pointing in opposite directions. The picture would look like one of the two options in Figure 5.2 except that one of the lines would be pointing in the other direction. In this case we want to rule out the case on the left hand side of Figure 5.2. Consider the case of vertical slices first. If the right hand picture occurs, then we have two light points on the same vertical grid line which occupy adjacent unit integer squares. In terms of the proof of the Exclusion Lemma I, we would have two points of opposite type, opposite direction, and odd separation. This is impossible. See the remark following the proof of the Exclusion Principle I.

The horizontal case is more subtle. We will prove that two opposite pointing worldlines intersect on the vertical midline of a unit square within $P/2$ from the a horizontal boundary H of the square. (Compare the left hand side of Figure 5.1.) This forces H to intersect both worldlines, and again we get the right hand side option.

Our analysis of the particles in §4.1 shows that every horizontal particle has an instance which lies in the boundary of a block. Hence, the worldlines in a horizontal slice all intersect the left and right edges of the rectangle at integer points. Let k denote the difference between the y -coordinates of the points where the lines intersect the left edge, as shown in Figure 5.3. (We have $k = 5$ in the figure.)

If k is even then the two worldlines cross on a vertical unit integer segment. If k is odd, then the two worldlines cross on the vertical midline of a unit integer square, as desired. We rule out the even case. If k is even, then there are two light points, both lying at the edges of their respective blocks, whose difference is a vector of the form $(2\ell\omega, 0)$ for some even integer ℓ . These points are light points on on a horizontal lines of the same height, and hence they have the same sign. On the other hand, these points differ by a vector of the form $(2\ell\omega, 0)$. Being at the edge of their blocks, they lie on slanting lines of all types. In particular, they each lie on slanting lines of slope $-P$ whose y -intercepts differ by $2\ell\omega P = 4\ell p$, an even number. But then our two light points have the same direction. This contradicts the worldline picture, in which the points have opposite direction.

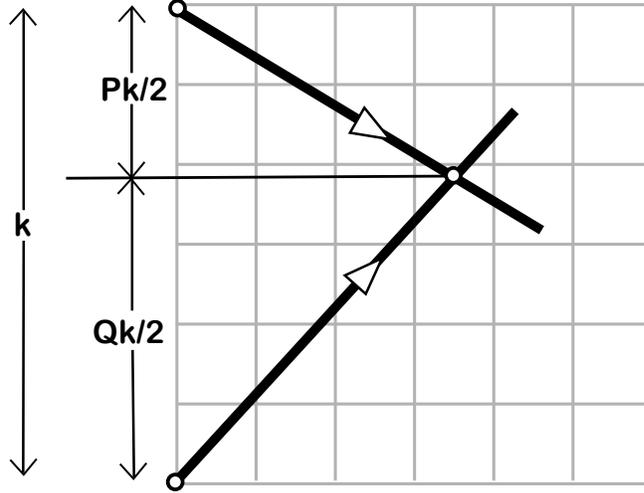


Figure 5.3: two worldlines intersecting

Now we know that k is odd as desired, and we come to the final point. The distance between the y -coordinate of the intersection of the two world lines and the top point on the left in Figure 5.3 is $Pk/2$. This comes from the fact that $P + Q = 2$ and that the two worldlines have slopes P and Q . We just need to show that

$$(Pk/2)_1 \in (-P/2, P/2). \quad (29)$$

Here $(y)_1$ denotes the difference between y and the integer nearest to y . Since $P = 2p/\omega$, Equation 29 is equivalent to

$$(2pk)_{2\omega} \in (-2p, 2p). \quad (30)$$

The light points ζ_1 and ζ_2 in the directed plaid model corresponding to the two left endpoints in Figure 5.3 differ by the vector $(k\hat{\tau}\omega, \pm\omega)$. Hence, the y -intercepts of the slanting lines of slope $-P$ containing ζ_1 and ζ_2 differ by $2k\hat{\tau}p \pm \omega$. These points have the same sign and mass in $(0, 2p)$. Given the definition of sign and mass, this forces

$$(2p \times (2pk\hat{\tau} \pm \omega))_{2\omega} \in (-2p, 2p).$$

But

$$(2p \times (2pk\hat{\tau} \pm \omega))_{2\omega} = (2p \times 2k\hat{\tau})_{2\omega} = (2pk \times 2p\hat{\tau})_{2\omega} = (2pk)_{2\omega}.$$

The last equality comes from the fact that $2p\hat{\tau}a \equiv 1 \pmod{\omega}$. This completes the proof.

6 Connection to the Truchet Tile System

6.1 Truchet Tilings

In this section we describe the Truchet tilings in $[\mathbf{H}]$ with $\alpha = \beta$, where $\alpha = p'/q' \in (0, 1/2]$. We insist that p' is odd and q' is even. We will see that every such tiling appears infinitely often as a pixellated spacetime diagram. Figure 6.1 shows the two Truchet tiles having “slope” $+1$ and -1 .

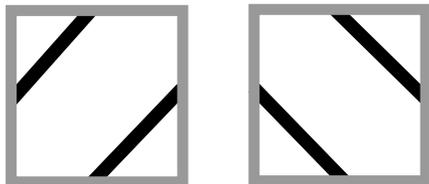


Figure 6.1: The truchet tiles.

The tilings in $[\mathbf{H}]$ corresponding to the parameters α, β are described at the beginning of §3 of $[\mathbf{H}]$. In general, the definition of the tiling depends on some choice of offset $(x, y) \in \mathbf{R}^2$, but when α and β are rational, all offsets give the same tiling up to translation. We take $\alpha = \beta$ and $x = y = -\alpha/2$. The functions ω and η in $[\mathbf{H}]$ are the same function.

Following $[\mathbf{H}]$, but using our notation, we define $\eta : \mathbf{Z} \rightarrow \{-1, 1\}$ to be the sign of

$$(m\alpha - \alpha/2)_1. \quad (31)$$

In other words, we take the sign of the representative of $m\alpha - \alpha/2 \bmod 1$ in $(-1/2, 1/2)$. This quantity is never 0, and the sequence is periodic with period q' . We call this sequence the *Truchet sequence*. We call the finite subsequence $\eta_1, \dots, \eta_{q'}$ the *first period*.

Here are some examples.

- When $\alpha = 1/2$, the sequence $\{\eta_m\}$ is given by $+ -$ repeating
- When $\alpha = 1/4$ the sequence is $+ + - -$ repeating.
- When $\alpha = 3/8$ the sequence is $+ - - + - + + -$ repeating.

We define a tiling of the plane in which the Truchet tiles are centered at integer points (m, n) and the slope of the Truchet tile is the same sign as $\eta_m \eta_n$. Figure 6.2 shows one period of the periodic Truchet tiling produced by these parameters.

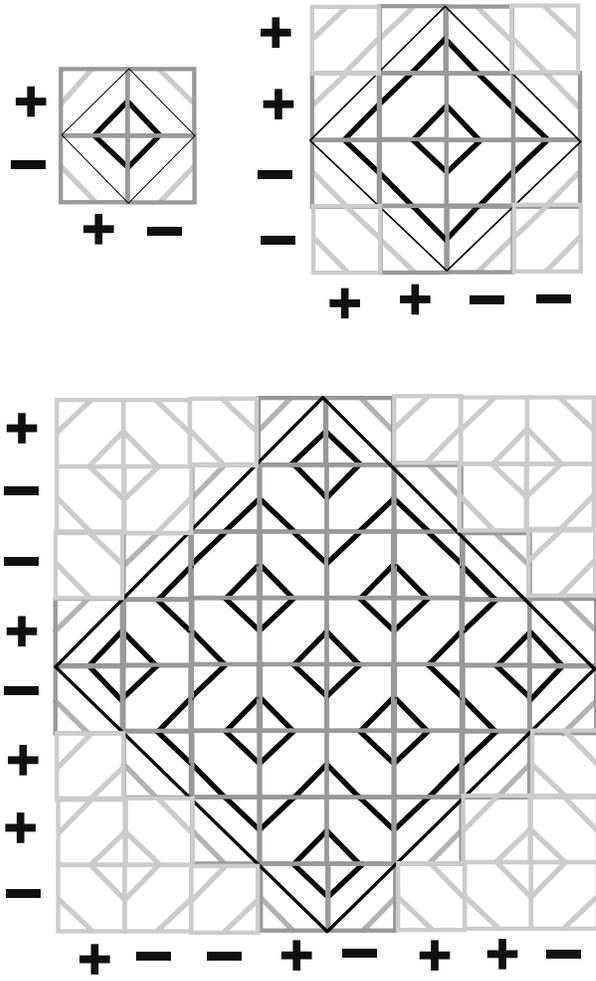


Figure 6.2: Truchet tile patterns for $\alpha = 1/2, 1/4, 3/8$.

The first period of the Truchet sequence has the form $(+A)(-A)$ where A is a palindrome. The first step in proving this is to write $f(m) = m\alpha - \alpha/2$ and observe that

$$f(m) + f(q + 1 - m) = p, \quad f(q/2 + m) - f(m) = p/2. \quad (32)$$

Given the structure of our sequence, we can always isolate a diamond shaped fundamental domain whose boundary is an impenetrable barrier disjoint from all Truchet paths. Hence, all Truchet paths are loops, and the picture always looks qualitatively like Figure 6.2.

6.2 The Truchet Comparison Theorem

Now for the moment of truth: We can interpret the Truchet tile systems as the output of curve following dynamics. We direct the vertical line $x = m$ upward if $\eta_m = 1$ and downward if $\eta_m = -1$. We direct the horizontal line $y = n$ rightward if $\eta_n = 1$ and leftward if $\eta_n = -1$. When we do the curve turning process and then omit the vertices which go through the centers of the squares, we recover the Truchet tiling. This is a purely local result, and easy to see from a single example. This fact is the basis for our main result below.

Let $\Gamma : \mathbf{R} \rightarrow \mathbf{R}$ denote the infinite dihedral group generated by the maps $x \rightarrow x - 1$ and $x \rightarrow 1 - x$. A fundamental domain for the action of Γ is the interval $[0, 1/2)$. Given any $x \in \mathbf{R}$, let $D(x)$ denote the representative of Γx in the fundamental domain. The map D plays an important role in [H].

Given a parameter p/q we define

$$\alpha(p/q) = D(P^{-1}) = D\left(\frac{p+q}{2p}\right). \quad (33)$$

For instance,

$$\alpha(4/9) = D(13/8) = 3/8, \quad \alpha(4/15) = D(19/8) = 3/8.$$

We chose this example because we have already drawn the Truchet pattern for $\alpha = 3/8$. Figure 6.2 shows half of the pixellated spacetime diagram for the vertical slice of capacity $8 = 2 \times 4$.

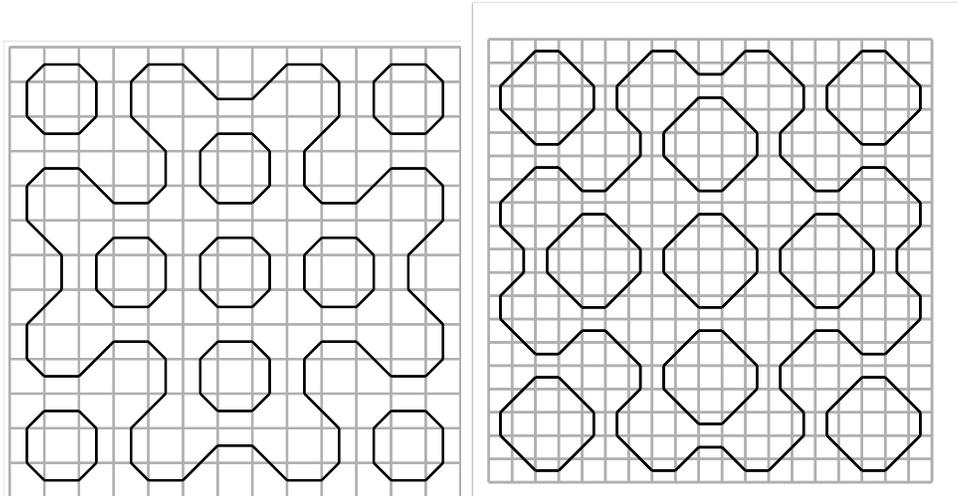


Figure 6.3: The vertical capacity 8 diagrams the parameter 4/9 and 4/15.

Figures 6.2 and 6.3 illustrate our main result. Notice the similarity between the Truchet pattern for $3/8 = \alpha(4/9) = \alpha(4/15)$ and the capacity 8 spacetime diagrams for 4/9 and 4/15.

Before we state our main result, we need one more notion. Within a Truchet tiling, we say that two curves *kiss* if they intersect the same Truchet tile. Within a pixellated spacetime diagram, we say that two polygons *kiss* if the corresponding orbits of the strict curve turning process meet at a vertex. The strict process traces over the curves exactly. Here is our main result.

Theorem 6.1 (Truchet Comparison) *Let $p/q \in (0, 1)$ be an even rational parameter. Then the union of curves in the pixellated spacetime diagram of capacity $2p$ for the parameter p/q is isotopic to the union of curves in two consecutive fundamental domains for the Truchet tiling associated to the parameter $\alpha(p/q)$. The isotopy preserves the kissing relation.*

Remarks:

- (i) Looking at Figure 6.3, we see that the figures have 8-fold dihedral symmetry. In the vertical case this always seems to be true, though our proof will only reveal an 8-fold combinatorial symmetry.
- (ii) As is pointed out in [H], the polygons in the Truchet pattern can be 2-colored so that kissing polygons have opposite colors. When the curves are equipped with the coloring, the fundamental domain needs to be twice as large. The same remarks apply to the kissing polygons in the spacetime diagrams.

Before proving the Truchet Comparison Theorem, we derive a corollary. Note, in Figure 6.3, that the largest polygon in each figure completely fills the square in which it is drawn. This is true in general, and is a consequence of the Truchet Comparison Theorem, because the corresponding polygon in the Truchet system always kisses its translates on all 4 sides. In particular, this largest polygon intersects every integer YZ slice in the undirected spacetime model except the mirrored boundaries.

Which surface could this largest polygon in the spacetime diagram belong to? Since the polygon intersects the ZY slices of capacity 2, and only the fundamental surface intersects this slice, we see that our polygon must lie on the fundamental surface. Hence, the fundamental surface intersects every non-boundary YZ slice in the undirected spacetime model. At the same time, our polygon intersects every integer XY slice except for one, the slice

corresponding to the bottom/top of the diagram. This slice turns out to be the block $B_{(\omega-1)/2}$, the one which has 4 fold dihedral symmetry. We call this XY slice the *symmetric slice*.

The same thing holds in the horizontal case, and this shows that the fundamental surface intersects every nonboundary XZ integer slice as well. Summarizing the argument above, we have the following result.

Theorem 6.2 (Filling) *For every even rational parameter, the fundamental plaid surface intersects every non-boundary integer coordinate slice in the undirected spacetime plaid model except perhaps for the symmetric slice.*

Remark: It can happen that the fundamental surface misses the symmetric slice. Figure 4.8 illustrates this for the parameter $1/2$. More generally, this happens for all parameters of the form $1/q$ and $1 - 1/q$. However, it seems that usually the fundamental surface does intersect the symmetric slice. This situation corresponds to the big polygon in the horizontal slice of capacity $2p$ slopping over the top and bottom of the square when it is drawn in the most symmetric way. Figure 6.4 shows this in action.

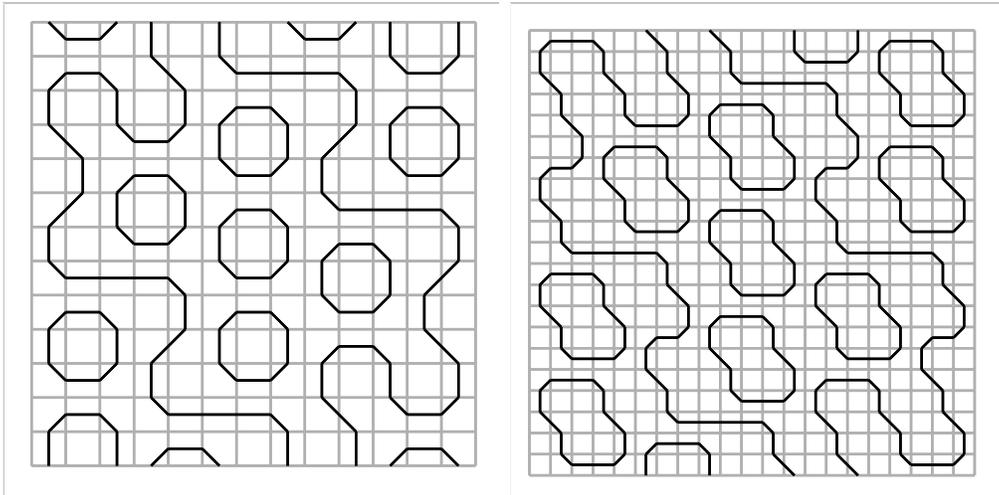


Figure 6.4: The vertical capacity 8 diagrams the parameter $4/9$ and $4/15$.

The rest of the chapter is devoted to the proof of the Truchet Comparison Theorem.

6.3 A Result from Elementary Number Theory

The Truchet Comparison Theorem relies on a curious result from elementary number theory. Let p/q be an even rational parameter. As usual, let $\omega = p+q$ and $\hat{\tau}$ be such that $2p\hat{\tau} \equiv 1 \pmod{\omega}$. We will carry along an example to help explain the result. In the example, $p/q = 5/8$. This gives $\omega = 13$ and $\hat{\tau} = 4$. Consider the following sequences.

Red Sequence: Define $A_k = \pm 1$ according to the sign of the representative of $(\omega/2 - \omega k) \pmod{2p}$ that lies in $(-p, p)$. This is the red sequence. In the example, the sequence of representatives is $(3/2, 9/2, -5/2, 1/2, 7/2)$ and the red sequence is $++-++$.

Blue Sequence: Take the numbers $\tau, 3\tau, 5\tau, \dots, (2p-1)\tau \pmod{\omega}$ and sort them so that they appear in order. Call the resulting sequence $\{y_k\}$. We define $B_k = +1$ if y_k is even and $B_k = -1$ if y_k is odd. The blue sequence is $\{B_k\}$. In the example, the unsorted sequence is $4, 12, 7, 2, 10$. The sorted sequence is $2, 4, 7, 10, 12$. The blue sequence is $++-++$.

Lemma 6.3 *The red and blue sequences coincide.*

Proof: Let $\langle m \rangle_p$ denote the representative of $m \pmod{p}$ that lies in $\{1, \dots, p\}$. Let $\langle m \rangle_{2p}$ denote the representative of $m \pmod{2p}$ in $\{-p+1, \dots, p\}$.

Let $c_1, \dots, c_p \in \{1, \dots, p\}$ be such that

$$y_k \equiv \hat{\tau}(2c_k - 1) \pmod{\omega}, \quad k = 1, \dots, p. \quad (34)$$

We will derive the following formula:

$$c_k = py_k + \frac{\omega + 1}{2} - \omega k. \quad (35)$$

If we introduce the convention that the sign of 0 is -1 then A_k is always the sign of

$$\hat{c}_k = \left\langle \frac{\omega + 1}{2} - k\omega \right\rangle_{2p}. \quad (36)$$

From Equation 35 and the fact that $c_k \in \{1, \dots, p\}$ we see that when y_k is even (respectively odd) we have $\hat{c}_k = c_k$ (respectively $\hat{c}_k = c_k - p$.) Hence y_k is even if and only if $A_k = +1$. ♠

To finish the proof of Lemma 6.3 we establish Equation 35. For ease of notation let $\tau = \widehat{\tau}$. When $p = 1$ we have $c_1 = 1$ and $y_1 = (\omega + 1)/2$. So, the formula holds. Henceforth assume that $p \geq 2$.

We set $y_0 = 0$ and $c_0 = 0$ for notational convenience. Define

$$\delta_k = c_{k+1} - c_k, \quad \Delta_k = y_{k+1} - y_k. \quad (37)$$

Lemma 6.4 $|\Delta_i - \Delta_j| \leq 1$ for all i, j .

Proof: Suppose $\Delta_i \leq \Delta_j - 2$. Let

$$c'_{j+1} = c_j + \delta_i + \epsilon p.$$

Here $\epsilon \in \{-1, 0, 1\}$ is chosen so that $c'_{j+1} \in \{1, \dots, p\}$. This is always possible because all the c variables are in $\{1, \dots, p\}$. Since $2\tau p \equiv 1 \pmod{\omega}$,

$$y'_{j+1} \equiv \tau(2c'_{j+1} - 1) \equiv \tau(2c_j - 1) + \Delta_i + \epsilon = y_j + \Delta_i + \epsilon \pmod{\omega}.$$

But then $y_j \leq y'_{j+1} < y_{j+1}$. Since y_j and y_{j+1} are consecutive, $y'_{j+1} = y_j$. But then $\Delta_i = 1$, which means that $2\tau\delta_i \equiv 1 \pmod{\omega}$. But this forces $\delta_i \equiv p \pmod{\omega}$. If $i > 0$ this is impossible because $c_i \neq c_{i+1} \in \{1, \dots, p\}$. If $\Delta_0 = 1$ then $y_1 = 1$. But then $\tau(2c_1 - 1) \equiv \tau(2p) \pmod{\omega}$ for some $c_1 \in \{1, \dots, p\}$. This is also impossible. ♠

Lemma 6.5 $\omega/2 < p\Delta_k < 3\omega/2$ for all $k = 0, \dots, p - 1$.

Proof: It is convenient to set $y_{p+1} = \omega + 1$. Observe that for each y_i there is some other y_j such that $y_i + y_j = \omega + 1$. But then $\Delta_p = \Delta_0$. We also have

$$\omega + 1 = y_{p+1} = \sum_{k=0}^p \Delta_k. \quad (38)$$

Let $\underline{\Delta}$ be the min of Δ_j and let $\overline{\Delta}$ be the max. If $\underline{\Delta} = \overline{\Delta}$ then the common value is $(\omega + 1)/(p + 1)$ and the inequalities are obvious.

Otherwise, we have $\overline{\Delta} = \underline{\Delta} + 1$, by the previous result. For the upper bound, we have $\underline{\Delta} \leq \omega/(p + 1)$, and this gives

$$p\overline{\Delta} = p\underline{\Delta} + p < \frac{p}{p+1}(\omega) + p < 3\omega/2.$$

For the lower bound we have $\overline{\Delta} \geq (\omega + 2)/(p + 1)$, and this gives

$$p\underline{\Delta} \geq \frac{p}{p+1}(\omega + 2) - p = \frac{p(q+1)}{p+1} > \omega/2.$$

A bit of calculus establishes the last inequality for $1 < p < q$. ♠

Let c_k^* denote the right hand side of Equation 35. We want to show that $c_k^* = c_k$. Using $2\tau \equiv 1 \pmod{\omega}$, we compute that $c_k^* \equiv c_k \pmod{\omega}$. Also, Lemma 6.5 tells us that

$$c_1^* = y_1 + \frac{\omega + 1}{2} - \omega = \Delta_0 + \frac{\omega + 1}{2} - \omega \in (0, \omega).$$

Since $c_1 \in (0, \omega)$ as well, we must have $c_1 = c_1^*$.

Suppose by induction that $c_k = c_k^*$. Then

$$c_{k+1}^* - c_{k+1} = (c_{k+1}^* - c_k^*) - (c_{k+1} - c_k) = (p\Delta_k - \omega) - \delta_k \in (-\omega, \omega).$$

The last entry of this equation comes from the fact that $|\delta_k| \leq p < \omega/2$, and from Lemma 6.5. Since $c_{k+1}^* \equiv c_{k+1} \pmod{\omega}$ and since $|c_{k+1}^* - c_{k+1}| < \omega$, we must have $c_{k+1}^* = c_{k+1}$. This completes the induction step. This establishes Equation 35.

6.4 Proof of the Truchet Comparison Theorem

We will give the proof for the spacetime diagrams corresponding to the slices of positive sign. The other case follows from symmetry.

There are $2p$ integers $0 < y_1 < \dots, y_{2p} < 2\omega$ such that the slanting lines having these y -intercepts have positive sign and mass less than $2p$. We call this the *low mass sequence*.

Lemma 6.6 *The sequence of directions of the low mass sequence coincides with the first period of the Truchet sequence, up to a global sign.*

Proof: This is an application of the fact that p/q is a good parameter, so that the red and blue associated sequences coincide. Let $\alpha = \alpha(p/q)$. Let $(+A)(-A)$ be the first period of the Truchet sequence associated to α . Let $\{A_k\}$ be the first half of this first period.

Let $\{y_k\}$ be the first half of the low mass sequence. This sequence is obtained exactly as the same-named sequence discussed in connection with the red sequence above. All the slanting lines in question have the same sign, and so their directions are determined by their parities. (See Lemma 1.3.) Therefore, up to a global sign, the sequence of directions of the first half of the low mass sequence is the red sequence associated to the parameter. To finish the proof we just have to recognize the first half of the first period of the Truchet sequence as the blue sequence associated to the parameter. The fact that the second halves agree follows from symmetry.

Let $[x]_s$ denote the sign of $(x)_s$. Let $\{A_k\}$ denote the first half of the first period of the Truchet sequence. The sequence associated to $\alpha + 1$ is $(+A)(-A)$ and the sequence associated to $1 - \alpha$ is $(-A)(+A)$. Hence, if $D(\alpha) = D(\alpha')$ then the associated Truchet sequences agree up to a global sign. Since $D(P^{-1}) = \alpha$, we see that

$$A_k = [P^{-1}k - P^{-1}/2]_1.$$

Multiplying through by $2p$ we get $\pm A_k = [\omega k - \omega/2]_{2p}$. The sign out in front is a global one. Hence the first half of the first period of the Truchet sequence agrees with the blue sequence. ♠

Consider the vertical case of the Truchet Comparison Theorem. To compare the pixellated spacetime diagrams to the Truchet tilings, we first extend the spacetime diagrams so that they fill the whole plane. We track the particles not just in the fundamental blocks but in the entire plane. This is to say that we consider the light points not just on the segment of a vertical grid line contained in a block, but on the entire line, and also we repeat the fundamental period endlessly. By symmetry, this extended picture is simply the orbifold universal cover of the original diagram. That is, we extend in the vertical direction by translations and in the horizontal direction by reflections. Figure 6.5 shows how we have reflected the original spacetime diagram in Figure 5.1 across one of its mirrored boundaries. Note how the directions match.

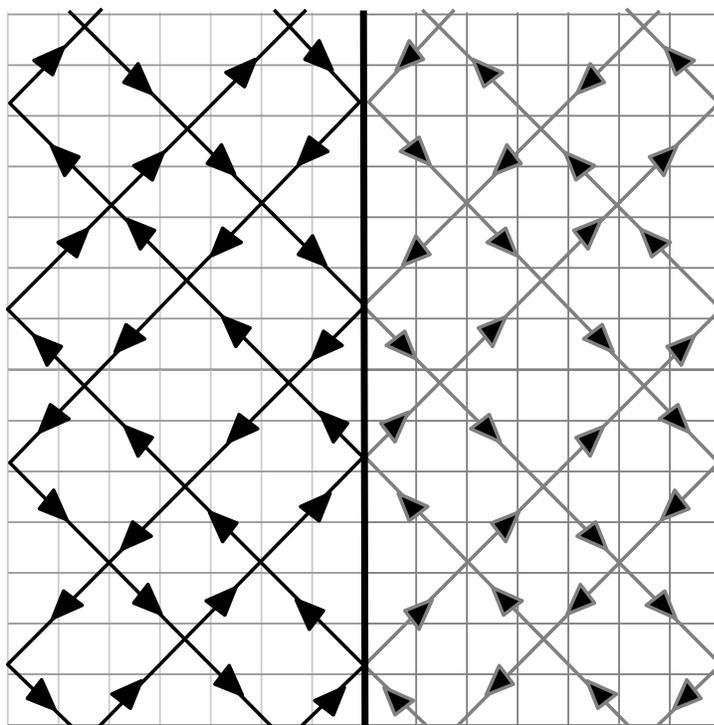


Figure 6.5: Two consecutive fundamental domains

Once we make the extension, the worldlines naturally fit together to make an infinite grid of directed lines of slope ± 1 . The lines of slope $+1$ intersect any given horizontal line in our diagram in the same positions that the slanting lines of slope $-P$ intersect the corresponding vertical grid line. The lines of slope -1 intersect the same horizontal line at the points where the slanting lines of slope $-Q$ intersect the corresponding vertical lines.

Note that the slanting lines of a given type intersect any vertical line in the same order as their y -intercepts. Indeed, the pattern of intersection in any vertical line is the same, up to translation, as the pattern of intersection on the y -axis. Therefore, assuming the detail above, the sequence of directions associated to the lines of slope $+1$ in the extended spacetime diagram is the same as the Truchet sequence, up to a global sign change. The same goes for the lines of slope -1 .

Given that the Truchet tiling can also be described in terms of curve turning, we see that the extended pixellated spacetime diagram is isotopic to the Truchet tiling in a way which preserves the kissing relation. The only difference is that we have turned the picture 45 degrees and also adjusted the

spacing between the lines. (Unlike in the Truchet tiling, the spacing is not always the same between different pairs of lines in the spacetime diagram.) Moreover, the original spacetime diagram contains 2 consecutive fundamental domains for the extended tiling. This completes the proof in the vertical case (modulo the detail), but below we will further analyze the symmetry of the situation.

The proof in the horizontal case is essentially the same. We just describe the differences. The first difference is that the lines in the extended spacetime model have slopes P and $-Q$ rather than 1 and -1 . This doesn't make any difference from a combinatorial point of view. Second, while the order in which the slanting lines intersect a given horizontal line is the same as the y -intercepts, the two sets are not translation equivalent. They are related by an affine transformation. This does not make any difference. This completes the proof of the Truchet Comparison Theorem

Remark: Our argument incidentally proves that the horizontal and vertical pixellated spacetime diagrams of capacity κ are combinatorially isomorphic to each other for all $\kappa \leq 2p$. The point is that both diagrams come from the curve turning construction, and in both slices we see the same combinatorial structure to the directions of the lines involved.

6.5 Symmetry of the Horizontal Diagrams

Now we discuss the extra symmetry. In the horizontal case, reflection in the *vertical* midline of the symmetric block $B_{(\omega-1)/2}$ permutes all the horizontal segments comprising the horizontal slices of the spacetime diagram, and also permutes the particles on these lines, preserving the directions. In the spacetime diagram this symmetry manifests itself as a rotation about the center point of the horizontal segment corresponding to the symmetric block.

Note furthermore that the spacetime diagram always has 2-fold translation symmetry. Translation by $(0, \omega)$ preserves all lines and reverses the directions. Combining the rotational and translation symmetry, we see that actually there are 4 points of rotational symmetry in the diagram, evenly spaced along the vertical midline. If we display the spacetime diagram so that the bottom line corresponds to the symmetric block, and we just show the bottom half, one of the centers of rotation is the center of the picture. This is why the examples in Figure 6.4 have 2-fold rotational symmetry.

Remark: This analysis works for all the horizontal spacetime diagrams. They all have this rotational symmetry.

The big polygon in the bottom half of the pixellated spacetime diagram – the one discussed in connection with the Filling Theorem – kisses the big polygon in the top half, because this is what we know happens in the Truchet tiling. But then the big polygon misses at most one horizontal line, as we have mentioned above. The symmetry we have just discussed shows that the missed horizontal segment, assuming it exists, must correspond to the symmetric block. Again, as we remarked above, this happens rarely.

6.6 Symmetry of the Vertical Diagrams

Reflection in the horizontal midline of any block permutes the light points and reverses the directions. This symmetry manifests itself as a reflection symmetry in the vertical midline of the spacetime diagram. This analysis works for any capacity. All these diagrams have vertical bilateral symmetry.

We have already remarked in the proof of Lemma 6.5 that for each y_i there is some y_j such that $y_i + y_j = \omega + 1$. Since these numbers are sorted, we must have $j = p + 1 - i$. In short,

$$y_k + y_{p+1-k} = \omega + 1, \quad k = 1, \dots, p. \quad (39)$$

This means that $(0, y)$ has positive sign and mass less than $2p$ if and only if $(0, \omega + 1 - y)$ does. At the same time, the points $(0, y_p)$ and $(0, \omega - y_p)$ have the same mass and opposite sign. This means that the points of mass less than $2p$ are *paired* in the sense that each point $(0, y)$ of mass less than $2p$ has a partner point $(0, y \pm 1)$ which has opposite sign and mass less than $2p$.

In every block, there are two vertical lines of capacity $2p$. Each slanting line produces a light point on one of the lines and the partner slanting line produces a light point on the other one. Relatively speaking, the two light points are in the same vertical integer segment. They are not actually in the same unit integer segment, however, because the one point is on the one line and the other point is on the other.

This structure has a special implication in the symmetric block. Reflection in the vertical midline of the symmetric block permutes the light points. Hence, in the symmetric block, the light points on each vertical segment of capacity $2p$ are paired. There are either 0 or 2 in each vertical unit integer

segment. In terms of the spacetime diagram, this means that all the lines of slope $+1$ *quasi-meet* all the all the lines of slope -1 in the corresponding horizontal slice: Both lines deposit light points in the same intervals. Take a hard look at the bottom of Figure 5.1. The lines do not meet precisely on the bottom line, but they meet up to one unit, and inside the same unit integer segments.

The bilateral symmetry of the diagram and the quasi-meeting property of the lines implies that reflection in the horizontal slice corresponding to the symmetric block is an approximate symmetry of the diagram itself and an exact symmetry if the pixellated diagram. The pixellated diagram has vertical and horizontal bilateral symmetry. Hence it also has the same points of reflection symmetry as in the horizontal case.

The same analysis as in the horizontal case now shows that the big polygon in each half of the spacetime diagram of capacity $2p$ can only miss the horizontal line corresponding to the symmetric block. This time, the quasi-meeting property implies that the big polygon does actually miss the segment we are talking about.

Remark: The pixellated vertical spacetime diagrams of all capacities seem to exhibit the same reflection symmetry mentioned above. I think that similar arguments as the ones given above would establish this. What is harder to understand is that all these diagrams seem to exhibit an exact 4 fold rotational symmetry. For the capacity $2p$ slice, the combinatorial 4-fold symmetry follows from the Truchet Comparison Theorem, thanks to the exact 4-fold rotational symmetry of the Truchet pattern.

Part II

The Plaid Master Picture Theorem

7 The Plaid Master Picture Theorem

7.1 The Spaces

Define

$$\widehat{X} = \mathbf{R}^3 \times [0, 1]. \quad (40)$$

The coordinates on \widehat{X} are given by (x, z, y, P) . We think of

$$P = \frac{2A}{1+A}, \quad A = p/q, \quad (41)$$

but P is allowed to take on any real value in $[0, 1]$. Define the following affine transformations of \widehat{X} .

- $T_X(x, y, z, P) = (x + 2, y + P, z + P, P)$.
- $T_Y(x, y, z, P) = (x, y + 2, z, P)$;
- $T_Z(x, y, z, P) = (x, y, z + 2, P)$;

Define two abelian groups of affine transformations:

$$\widehat{\Lambda}_1 = \langle T_X, T_Y, T_Z \rangle, \quad \widehat{\Lambda}_2 = \langle T_X^2, T_Y, T_Z \rangle. \quad (42)$$

Finally define

$$X_k = \widehat{X} / \Lambda_k, \quad k = 1, 2. \quad (43)$$

The space X_2 is a double cover of X_1 . Both spaces should be considered flat affine manifolds - i.e. manifolds whose overlap functions are restrictions of affine transformations. All the affine transformations in sight preserve the slices $\mathbf{R}^3 \times \{P\}$ and act as translations on these slices. This X_1 and X_2 are fibered by 3-dimensional Euclidean tori. We denote these tori by $X_1(P)$ and $X_2(P)$ respectively.

Remark: The reason why we keep track of two spaces is that a suitable partition of X_1 into polytopes determines the unoriented polygons in the plaid model whereas a suitable partition of X_2 determines the oriented polygons in the plaid model. Often it is easier to do computations with X_1 even though X_2 contains more information.

7.2 The Partition

Here we describe a partition of the space X_1 into convex integral polytopes. Taking the double cover, we get a partition of X_2 . Lifting to \widehat{X} , we get a Λ_1 -invariant partition of \widehat{X} into infinitely many polytopes. When we add labels, this infinite partition is only Λ_2 -invariant. Our computer program lets the user see the partition in an interactive way.

The rectangular solid

$$[-1, 1]^3 \times [0, 1]. \quad (44)$$

serves as a fundamental domain for the action of Λ_1 on \widehat{X} . We think of this space as a fiber bundle over the (x, P) plane. The base space B is the rectangle $[-1, 1] \times [0, 1]$. The space B has a partition into 3 triangles, as shown in Figure 7.1.

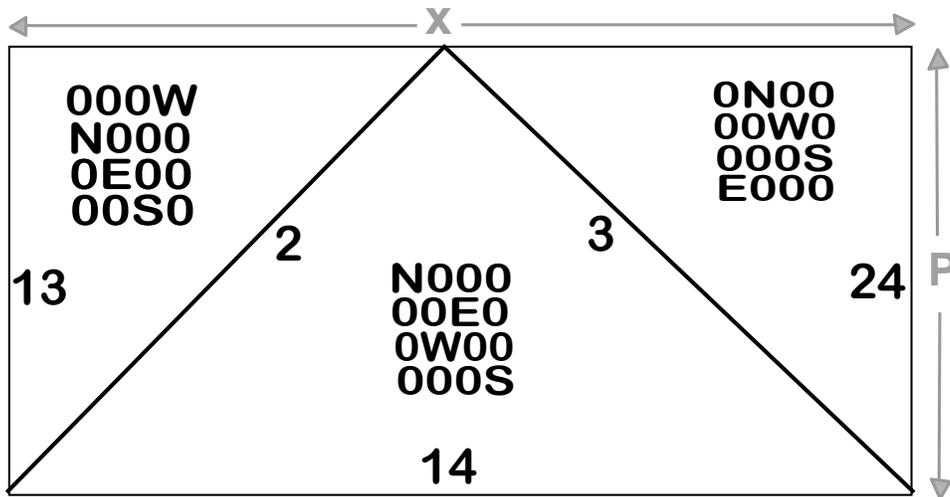


Figure 7.1: The (x, P) base space.

Each triangular region has been assigned a 4×4 matrix which we will explain momentarily. The slices above each open triangle in the partition of B consists of a 4×4 grid of rectangles. For of the rectangles, corresponding to the polytopes labeled by the emptyset are squares. (In Figure 7.2 we replace the emptyset by the symbols $\{S, W, N, E\}$ because this makes the rest of the labeling pattern more clear. The pattern of the squares is indicated by the nonzero entries in the matrices. For instance, the grid shown in Figure 7.2 corresponds to a fiber over the left triangle. We have also indicated the labels in the picture.

WN	WE	WS	W
N	NE	NS	NW
EN	E	ES	EW
SN	SE	S	SW

Figure 7.2: The checkerboard partition

The labels can be derived from the matrices in Figure 7.1. Every non-square rectangle in a column has the same first label and every rectangle in a row has the same second label.

The edge labels in Figure 7.1 encode degenerations in the fibers. Away from the singular edges of the base space, each polytope is a rectangle bundle over an open triangle. As one approaches the edges of the triangles in the base, the partition degenerates in that some of the rectangles shrink to line segments or points. The numbers on the edges in Figure 7.1 indicate which rows and columns degenerate. For instance, as we approach the edge labeled 2, the thickness of second column from the left tends to 0 and at the same time the thickness of the second row from the bottom tends to 0. The whole (unlabeled) picture is symmetric with respect to reflection in the line $y = x$.

From the description of these degenerations and from the fact that we know we are looking at slices of convex integral polytopes, one can actually reconstruct the entire partition from the labelings in Figure 7.1. Each polytope in Figure 7.2 comes with a label. For instance, the label of the polytope which intersects the fiber in the northwest corner is $\{W, N\}$.

Thus, each polytope in the partition of \widehat{X} either is either labeled by the

\emptyset symbol, or else by a pair $A \neq B$, where $A, B \in \{N, E, W, S\}$. Now we explain how we enhance the labeling so that each polytope not labeled by \emptyset is labeled by an *ordered pair* (A, B) . Our idea is to explicitly give the ordering on a list of 10 polytopes (below) and then use the action of Λ_1 to propagate the ordering to the whole partition. The group Λ_2 preserves the ordered labeling and elements of $\Lambda_1 - \Lambda_2$ reverse the ordering. We found the ordered labeling (which turns out to be canonical) by trial and error.

Here we list the vertices and ordered labels of 10 polytopes. All the polytopes in the partition of \widehat{X} can be obtained from these using the group generated by Λ_1 and the following two additional elements:

- **Negation:** The map $(x, y, z, P) \rightarrow (-x, -y, -z, P)$ preserves the partition and changes the labels as follows: N and S are swapped and E and W are swapped.
- **Flipping:** The map $(x, y, z, P) \rightarrow (x, z, y, P)$ preserves the partition and changes the labels as follows: N and S are swapped, and the order of the label is reversed.

The last polytope listed has 12 vertices, and the listing is spread over 2 lines.

$$\begin{array}{ccccc}
 \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} & & (W, E) \\
 \\
 \begin{bmatrix} +1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ +1 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ -1 \\ +1 \end{bmatrix} & & (E, S) \\
 \\
 \begin{bmatrix} +1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ -1 \\ -1 \\ +1 \end{bmatrix} & & (E, N) \\
 \\
 \begin{bmatrix} -1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ +1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ +1 \\ +1 \end{bmatrix} & (W, S)
 \end{array}$$

$$\begin{array}{cccccccc}
\begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ 0 \\ +1 \end{bmatrix} & (S, W) \\
\begin{bmatrix} -1 \\ +1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} -1 \\ 0 \\ +1 \\ +1 \end{bmatrix} & \begin{bmatrix} -1 \\ +1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} -1 \\ +1 \\ +1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ +1 \\ +1 \\ +1 \end{bmatrix} & \begin{bmatrix} -2 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \emptyset \\
\begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ -1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ -1 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ 0 \\ +1 \end{bmatrix} & (S, N) \\
\begin{bmatrix} +1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} 2 \\ +1 \\ 0 \\ +1 \end{bmatrix} & (E, W) \\
\begin{bmatrix} -1 \\ -1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ -1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ -1 \\ +1 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ +1 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ 0 \\ +1 \\ +1 \end{bmatrix} & \begin{bmatrix} +1 \\ -1 \\ +1 \\ 0 \end{bmatrix} & \emptyset \\
\begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ +1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ +1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} +1 \\ +1 \\ +1 \\ 0 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ +1 \end{bmatrix} & & & \\
\begin{bmatrix} -1 \\ 0 \\ -1 \\ +1 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} -1 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} & \begin{bmatrix} -3 \\ -1 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} -2 \\ -1 \\ -1 \\ +1 \end{bmatrix} & & & \emptyset
\end{array}$$

7.3 The Map

The *plaid grid* is defined to be the set G of centers of integer unit squares. For each parameter $A \in (0, 1)$, there is a linear map $\Phi_A : G \rightarrow \widehat{X}$, given by

$$\Phi_A(x, y) = (2Px + 2y, 2Px, 2Px + 2Py, P), \quad P = \frac{2A}{1+A}. \quad (45)$$

In this section we fix A and $P = 2A/(1+A)$ and we write $\Phi = \Phi_A$. Also, we restrict the action of Λ_1 to $\mathbf{R}^3 \times \{P\}$ and we identify this slice with \mathbf{R}^3 . Thus, we think of Λ_1 as a lattice of translations acting on \mathbf{R}^3 .

Lemma 7.1 *Let L_1 be the lattice of symmetries of the undirected plaid model and let L_2 be the lattice of symmetries of the directed plaid model. If we interpret Λ_j as a lattice of translation vectors, then $\Phi(L_j) \subset \Lambda_j$ for $j = 1, 2$*

Proof: We will prove this for $j = 1$. The proof for $j = 2$ is similar. Recall that $A = p/q$ and $P = 2A/(1+A)$ and $\omega = p+q$ and that L is generated by $(\omega^2, 0)$ and $(0, \infty)$. Since Φ is a linear map, it suffices to prove that $\Phi(\omega^2, 0)$ and $\Phi(0, \omega)$ both belong to Λ_1 . We compute

$$\Phi(\omega^2, 0) = (2P\omega^2, 2P\omega^2, 2P\omega^2) =$$

$$P\omega^2(2, P, P) + (0, 2P\omega^2 - P^2\omega^2, 0) + (0, 0, 2P\omega^2 - P^2\omega^2).$$

Note that $P = 2p/\omega$, so that $P\omega^2$ and $P^2\omega^2$ are both even integers. Hence, the last two vectors above have the form $(0, k, 0)$ and $(0, 0, k)$ for some even integer k . All the vectors on the second line of our equation belong to Λ_1 .

We compute

$$\Phi(0, \omega) = (2\omega, 0, 2P\omega) = \omega(2, P, P) + (0, -\omega P, 0) + (0, 0, 0P\omega).$$

Again, the last two vectors have the form $(0, -k, 0)$ and $(0, 0, k)$ for some even integer k . So, all the vectors on the right hand side belong to Λ_1 . ♠

Recall that \mathbf{Z}_0 and \mathbf{Z}_1 respectively denote the set of even and odd integers, we define

$$\mathcal{X} = \left(\frac{\mathbf{Z}_1}{\omega}, \frac{\mathbf{Z}_0}{\omega}, \frac{\mathbf{Z}_0}{\omega} \right). \quad (46)$$

Note that the lattice $\Lambda_1(P)$ consists entirely of vectors in $(\mathbf{Z}_0/\omega)^3$. Hence Λ_1 preserves \mathcal{X} . The quotient \mathcal{X}/Λ_1 is a finite set of points with ω^3 members. The quotient G/L is also a finite set of points with ω^3 members.

Lemma 7.2 Φ induces a bijection between G/L_j and \mathcal{X}/Λ_j for $j = 1, 2$.

Proof: We prove this for $j = 1$. The case of $j = 2$ is similar, and indeed follows from the case $j = 1$ and symmetry. Looking at the formula in Equation 45, we see that $\Phi(c) \in \mathcal{X}$. Hence, by Lemma 7.1, the map Φ induces a map from G/L to \mathcal{X}/Λ_1 . In view of the fact that both sets have the same number of points, we just have to show that the induced map is an injection.

Suppose that $\Phi(c_1) = \Phi(c_2)$. We write

$$c_2 - c_1 = (x, y) = \left(\frac{m}{\omega}, \frac{n}{\omega} \right), \quad m, n \in \mathbf{Z}. \quad (47)$$

We have

$$\Phi(c_2) - \Phi(c_1) = \left(\frac{2pm}{\omega^2}, \frac{2pm}{\omega^2} - \frac{pn}{\omega^2}, \frac{2pm}{\omega^2} - \frac{Pn}{\omega^2} \right) \pmod{\Lambda}.$$

In order for the first coordinate to vanish mod Λ , we must have $m \equiv 0 \pmod{\omega^2}$. For the second and third coordinates to vanish mod Λ , we must have $n \equiv 0 \pmod{\omega}$. But then $c_2 - c_1 \in L_1$. ♠

Now we come to the main point in this section.

Lemma 7.3 (Definedness) $\Phi(G)$ is disjoint from the boundaries of the polytopes in the partition of \widehat{X} .

Proof: In view of our previous results, we just need to know that \mathcal{X} does not hit the boundaries of polytopes in our partition. We forget the last coordinate and work in \mathbf{R}^3 . Say that a *special plane* is a plane of the form $\{T\} \times \mathbf{R}^2$ with $T \in \mathbf{Z}_1/\omega$. By definition, \mathcal{X} is contained in the union of special planes.

We check that each special plane intersects the walls of the partition in lines of the form $x = u$ and $y = u$. When $T = -1$, the values of u are $\{-1, 1 + P, 1\}$, all of which belong to \mathbf{Z}_1/ω . As the value of T changes by $2/\omega$, the offsets for the wall-fiber intersections change by $\pm 2/\omega$. Hence, we always have $u \in \mathbf{Z}_1/\omega$. But then \mathcal{X} does not hit any of these lines.

The proof is done, but we want to elaborate on the picture. If we place a cube of side length $2/\omega$ around each point of \mathcal{X} , then these cubes tile \mathbf{R}^3 . Moreover, these cubes intersect each special plane in an $\omega \times \omega$ grid. The image $\Phi(G)$ intersects this grid at the centers of the squares whereas the walls of the partition intersect the grid in line segments extending the edges of the squares. ♠

7.4 Three Dimensional Interpretation

We can think of Φ as a map from the set of centers horizontal unit integer squares in the cubical plaid model. In these coordinates, the vector $(0, 0, 1)$ corresponds to $(-\hat{\tau}, \pm\omega)$. The sign depends on the parity of the z -coordinate of the horizontal slice at which the vector $(0, 0, 1)$ is anchored.

The set of points in \mathbf{R}^3 corresponding to points in the plaid grid have the form (x, y, z) where x, y are half integers and z is an integer. In this section, we will give a formula for $\Phi(x_2, y_2, z_2) - \Phi(x_1, y_1, z_1)$ where (x_1, y_1, z_1) and (x_2, y_2, z_2) are two points of this kind.

Lemma 7.4 *Let $\Phi : \mathbf{Z}^3 \rightarrow \mathbf{R}^3/\Lambda_2$ be the map given by*

$$\Phi(x, y, z) = x(2P, 2P, 2P) + y(2, 0, 2P) + z(-2, 0, 0) \bmod \Lambda_2. \quad (48)$$

This formula computes

$$\Phi(x_1, y_1, z_1) - \Phi(x_2, y_2, z_2)$$

where $(x, y, z) = (x_2, y_2, z_2) - (x_1, y_1, z_1)$ in the cubical coordinates.

Proof: We have

$$\Phi(1, 0) = (2P, 2P, 2P), \quad \Phi(0, 1) = (2, 0, 2P). \quad (49)$$

We also have

$$\begin{aligned} \Phi(-\hat{\tau}\omega, 0) &= (2P\hat{\tau}\omega, 2P\hat{\tau}\omega, 2P\hat{\tau}\omega) = (4p\hat{\tau}, 4p\hat{\tau}, 4p\hat{\tau}) \equiv_1 \\ &(4p\hat{\tau}, 0, 0) \equiv_2 (0, -2Pp\hat{\tau}, -2Pp\hat{\tau}) \equiv_3 (0, P, P) \bmod \Lambda_2. \end{aligned} \quad (50)$$

Equivalence 1 comes from the fact that $4p\hat{\tau} \in 2\mathbf{Z}$. Equivalence 2 comes from subtracting multiples of $(2, P, P)$. Equivalence 3 comes from the fact that

$$-2Pp\hat{\tau} - P = P(-2p\hat{\tau} - 1) = PK\omega = 2pK \in 2\mathbf{Z}.$$

Next, we compute

$$\begin{aligned} \Phi(0, \omega) &= (2\omega, 0, 2P\omega) \equiv (2\omega, 0, 0) = \\ &(2(\omega - 1) + 2, 0, 0) \equiv (2, -P\omega + P, -P\omega + P) \equiv (2, P, P) \bmod \Lambda_2. \end{aligned} \quad (51)$$

This combines with Equation 50 to give

$$\Phi(-\hat{\tau}, \pm\omega) \equiv (0, P, P) \pm (2, P, P) \equiv (-2, 0, 0) \bmod \Lambda_2. \quad (52)$$

Putting all this together, we get Equation 48. ♠

7.5 The Main Result

Once again we fix an even rational parameter A and set $P = 2A/(1+A)$. We use the same notation conventions from the previous section. The partition of \widehat{X} by labeled polytopes descends to partition of X_2 .

The map $\Phi : G \rightarrow X_2$ determines a tiling of \mathbf{R}^2 , as follows. In the unit integer square centered at $c \in G$ we place...

- the empty tile if $\Phi(c)$ lies in a polytope labeled by \emptyset , or
- the tile with a directed segment pointing from the midpoint of the α edge to the midpoint of the β edge if $\Phi(c)$ lies in a polytope with label (α, β) .

We call this tiling the *directed tiling*. Given a unit square \square and an edge E of \square we say that the directed tiling *involves* (\square, E) if the connector in \square uses the midpoint of E . In this case, we call (\square, E) *out-pointing* if the connector points to E and *in-pointing* if the connector points from E . Here is our main result, which holds for every parameter.

Theorem 7.5 *For any unit integer square \square and any edge E of \square , the number light points assigned to E is odd if and only if the directed tiling involves (\square, E) . In this case, (\square, E) is in-pointing if and only if the transverse direction associated to the light point on E in §1.6 points into \square .*

In short, Theorem 7.5 says that Theorem 1.1 is true, that the tiling obtained by the plaid construction matches the tiling obtained from the PET, and that the transverse directions given in §1.6 match the transverse directions given by the PET. There is a complete isomorphism. As an immediate corollary, the method of assigning directions in §1.6 is tile consistent.

Now we rephrase Theorem 7.5 in terms of the Plaid Master Picture Theorem. Consider the map $f : G \rightarrow G$ that comes from simply following the arrows on the tiles. Given $c_0 \in G$, the new point $C_1 = f(c_0)$ is defined to be the center of the tile into which the tile at c_0 points. For instance, the tile centered at c_0 is NE, then $c_1 = c_0 + (1, 0)$. In case the tile centered at c_0 is empty, we have $c_1 = c_0$.

Thanks to Theorem 7.5, there is a corresponding map $F : \widehat{X} \rightarrow \widehat{X}$ such that

$$F \circ \Phi = \Phi \circ f. \tag{53}$$

The map F has the following definition:

- F is the identity on polytopes labeled by \emptyset .
- $F(x, y, z) = (x, y, z) + (2, 0, 2P)$ on polytopes whose label ends in N .
- $F(x, y, z) = (x, y, z) - (2, 0, 2P)$ on polytopes whose label ends in S .
- $F(x, y, z) = (x, y, z) + (P, P, P)$ on polytopes whose label ends in E .
- $F(x, y, z) = (x, y, z) - (P, P, P)$ on polytopes whose label ends in W .

We can also interpret F as a map

$$F_2 : X_2 \rightarrow X_2 \tag{54}$$

because everything in sight is Λ_2 invariant. With this interpretation, (X_2, F_2) is the fibered integral affine PET. What we mean is the 4 dimensional system is piecewise affine, and that the PET preserves the P -slices and acts as an ordinary piecewise translation in each fiber. The 4 dimensional system is integral in the sense that all the polytopes in the partition have integer vertices. Thus, Theorem 7.5 implies the Plaid Master Picture Theorem from the introduction.

7.6 The Undirected Result

In Part 4 of the monograph, when we prove the Quasi-Isomorphism Theorem, we will simplify our calculations by using the space X_1 . Accordingly, we say about the undirected version of Theorem 7.5. The polytopes which partition X_1 are labeled by unordered pairs of the form $\{N, E\}$, etc. We cannot order these pairs in a canonical way because the action of elements in $\Lambda_1 - \Lambda_2$ reverses the orderings within our labels. Thus, some element of $\Lambda_1 - \Lambda_2$ maps a polytope labeled (N, E) to a polytope labeled (E, N) .

Nonetheless, we can use the unordered labels to define an undirected tiling in the same way that we defined the directed tiling. We just leave off the directions. Theorem 7.5 has the following immediate corollary.

Theorem 7.6 *For any unit integer square \square and any edge E of \square , the number light points assigned to E is odd if and only if the undirected tiling involves (\square, E) .*

Theorem 7.6 is good enough for the proof of the Quasi-Isomorphism Theorem.

8 The Images of Particles

8.1 Overview

The next three chapters are devoted to the proof of Theorem 7.5. As always, we fix a parameter $A = p/q$ and set $\omega = p + q$.

Each particle in the directed model breaks up into two *half-particles* in which the relative motion of the subset goes in one direction across the block. In the horizontal cases, one of the half-particles has $2p$ consecutive instances and moves east and the other half-particle has $2q$ instances and moves west. In the vertical case, one of the half-particles moves south and the other moves north. Both half-particles have ω instances in this case.

Let z_0, \dots, z_{n-1} be the successive instances of a vertical (respectively horizontal) half-particle and let c_0, \dots, c_{n-1} be the corresponding sets of centers of the unit squares containing these points on the west (respectively south) edge. We are interested in the image

$$Z = \bigcup_{i=0}^{n-1} \Phi(c_j) \subset X_1(P). \quad (55)$$

Our basic result is that *the images of half-particles under the classifying map travel in straight lines*. Here is the precise statement. Let (T, U_1, U_2) be coordinates on this fundamental domain. This chapter is devoted to proving the following result, which works for both light and dark particles.

Lemma 8.1 (Particle) *The following is true.*

1. *If Z corresponds vertical P particle, then modulo Λ_2 , the set Z lies in a segment with endpoints $(T, -1, U_2)$ and $(T, 1, U_2)$.*
2. *If Z corresponds vertical Q particle, then modulo Λ_2 , the set Z lies in a segment with endpoints $(T, U_1, -1)$ and $(T, U_1, 1)$.*
3. *If Z corresponds horizontal P particle, then modulo Λ_2 , the set Z lies in a segment with endpoints of the form of the form $(\pm 1, U_1, U_2) - (P, P, P)$ and $(\pm 1, U_1, U_2) + (P, P, P)$.*
4. *If Z corresponds horizontal Q particle, then modulo Λ_2 , the set Z is a segment with endpoints of the form of the form $(\pm 1, U_1, U_2) - (Q, 0, 0)$ and $(\pm 1, U_1, U_2) + (Q, 0, 0)$.*

8.2 The Vertical Case

If c_k and c_{k+1} are two successive instances of a vertical P half-particle then in terms of the 3 dimensional model $c_{k+1} - c_k = (0, 1, 1)$. Hence, by Equation 48,

$$\Phi(c_{k+1}) - \Phi(c_k) \equiv (0, 0, 2P) \pmod{\Lambda_2}.$$

This proves Statement 1.

In the vertical type Q case, we have $c_{k+1} - c_k = (0, -1, 1)$. Hence, by Equation 48,

$$\Phi(c_{k+1}) - \Phi(c_k) \equiv (-4, 0, -2P) \equiv (0, 2P, 0) \pmod{\Lambda_2}.$$

This proves Statement 2. .

8.3 The Horizontal P Case

Our first result vastly simplifies the amount of calculating we have to do.

Lemma 8.2 *If Statement 3 is true for one horizontal P particle then it is true for all of them.*

Proof: This is most easily seen in terms of the 3-dimensional interpretation of the map. All the horizontal P particles have instances of the form $(0, y, z)$ for $y, z \in \mathbf{N}$. By Equation 48, the sets corresponding to the two particles differ from each other by a translation of the form $(0, y', z')$ or $(2, y', z')$. ♠

We will compute using the (dark) half particle which has an instance at $(0, 0)$ in the plane. In terms of the 3-dimensional interpretation of the model, the successive instances z_0, \dots, z_{2p-1} are given by

$$z_k = (k/P, 0, k). \tag{56}$$

See (the gray part of) Figure 4.1.

Let $\{x\}$ denote the half integer nearest x . We have

$$\{x\} = x - [x] + 1/2, \tag{57}$$

where $[x]$ is the number in $(0, 1)$ representing $x \pmod{\mathbf{Z}}$. These equations are not defined when $x \in \mathbf{Z}$, but this situation does not arise for us.

The square center c_k corresponding to z_k is

$$(\{k/P\}, 1/2, k). \quad (58)$$

In the two dimensional model, c_0 has coordinates $(1/2, 1/2)$. Using the original definition of the map Φ we see that

$$\Phi(c_0) = (P - 1, 0, P). \quad (59)$$

Note that $c_k - c_0 \in \mathbf{Z}^3$, so we may use Equation 48 to compute. Equation 48 gives

$$\Phi(c_k) - \Phi(c_0) = (\mu_k, \mu_k, \mu_k) + (-2k, 0, 0) \bmod \Lambda_2, \quad (60)$$

where

$$\mu_k = 2P\{k/P\} - 2P\{0/P\} = 2k - 2P[k/P]. \quad (61)$$

Setting

$$\nu_k = -2P[k/P] \in [-2P, 0], \quad (62)$$

, we get

$$\Phi(c_k) - \Phi(c_0) \equiv (\nu_k, \nu_k, \nu_k) \bmod \Lambda_2, \quad (63)$$

Hence $\Phi(c_k)$ lies on the line segment with endpoints

$$(-1 + P, 0, P), \quad (-1 - P, -2P, -3P). \quad (64)$$

This proves Statement 3.

8.4 The Horizontal Q Case

The same argument as in Lemma 8.2 reduces this case to the study of a single horizontal Q particle. We will compute using the (dark) particle which (again) has an instances z_1, \dots, z_{2q} , where

$$z_k = (k/Q, 0, 2\omega - k) \quad (65)$$

Remark: We are abusing our notation and conventions to make the calculation as clean as possible. First, we are tracing out instances going “backwards in time”, so that in Figure 4.1 we would be tracing the black worldline from top to bottom. Second, we are allowing the third coordinate to lie in $[0, 2\omega]$ rather than in $[0, \omega]$, as we would want for the undirected model. However, for the purposes of verifying Statement 4 of the Particle Lemma, there is no

harm in replacing a point (a, b, c) with the point $(a, b, c \pm \omega)$.

The centers corresponding to our points are

$$c_k = (\{k/Q\}, 1/2, 2\omega - k), \quad k = 1, \dots, 2q. \quad (66)$$

It is convenient to add c_0 to the list of centers.

Before we do the calculation, we use the fact that $P + Q = 2$ to derive an alternate version of Equation 48.

$$\Phi(x, y, z) = x(-2Q, 0, 0) + y(2, 0, -2Q) + z(-2, 0, 0) \bmod \Lambda_2. \quad (67)$$

Again, this only works when $x, y, z \in \mathbf{Z}$.

We have $\Phi(c_0) = (P - 1, 0, P)$ as in the previous case. Using Equation 67, we have

$$\begin{aligned} \Phi(c_k) - \Phi(c_0) &= \left(-2Q(\{k/Q\} - \{0/Q\}), kQ, kQ \right) \equiv \\ &= (2Q[k/Q], 0, 0) \bmod \Lambda_2. \end{aligned} \quad (68)$$

From this equation we see that $\Phi(c_k)$ lies in the segment with endpoints

$$(1 - Q, 0, P) = (P - 1, 0, P), \quad (1 + Q, 0, P) = (P - 1 + 2Q, 0, P).$$

This proves Statement 4.

Remark: The same remarks about symmetric instances of the particle apply in this case as well.

9 Proof of the Main Result

9.1 Symmetric Instances

In the previous chapter we understood the images of half-particles under the map Φ . As in the previous chapter, we suppress the coordinate which records the parameter and we work in \mathbf{R}^3 . In this chapter we will prove the main result modulo two lemmas which we prove in subsequent chapters.

We say that a *symmetric instance* of a half-particle is an instance that lies in a unit integer square that has its midpoint on the horizontal midline of a block. Every half-particle has a symmetric instance. If z is a symmetric instance of vertical (respectively horizontal) particle, let c be the center of the unit integer square which has z on its west (respectively south) edge. In this chapter we will study the image of $\Phi(c)$. We call this the point the *character* of z . And we denote it by $\phi(z)$. Note that $\phi(z) = \Phi(c)$.

Remark: In making this definition we have arbitrarily chosen to favor the west and south edges over the east and north edges. Were we to make the switch, we'd get similar answers. We will arrange our proof so that we do not need to consider the east and north edges. The point is that every east or north edge of a square is also a west or south edge of an adjacent square.

Lemma 9.1 *Let z be a symmetric instance of a vertical particle. Then mod Λ_2 , the character of z has the form (T, U, U) .*

Proof: We have $c = (x, y)$ where $y = \pm\omega/2$. Note that $2P\omega = 2p \in 2\mathbf{Z}$. Hence, Equation 45 gives $\Phi(c) = (2Px, 2Px, 2Px) + (\pm\omega, 0, 0)$. ♠

Lemma 9.2 *Let z be a symmetric instance of a horizontal particle. Then mod Λ_2 , the character of z has the form $(\pm 1, U_1, U_2)$.*

Proof: This time we have $c = (x, y)$ where $x = k\omega/2$ for some integer k . Also, $2y$ is an odd integer. As in the previous result, $2Pk\omega/2 \in 2\mathbf{Z}$. But then, by Equation 45, the first coordinate of $\Phi(c)$ is an odd integer. Mod $4\mathbf{Z}$ we can take this coordinate to be ± 1 . ♠

9.2 Sharp Containers

We identify 4 kinds of vertical light points, namely

$$(\{P, Q\}, \{E, W\}).$$

For example, the light point of type (P,west) lie on slanting lines of slope $-P$ which are directed so that they point west. Likewise we identify 4 kinds of horizontal light points, namely

$$(\{P, Q\}, \{N, S\}).$$

These names have similar interpretations. We also classify the dark points, but we ignore the directions. Thus there are horizontal dark points of type P and Q, and vertical dark points of type P and Q. We say that a dark point has the same *undirected type* as a light point if, when we forget the direction, we get the same type. For instance, a vertical dark point of type P has the same type as the vertical light points of type (P,east) and (P,west).

We say that a set $\Sigma \subset \mathbf{R}^3$ is a *container* for a given type

1. The character of a light point with the given type always maps into $\Sigma \bmod \Lambda_2$.
2. The character of a dark point with the same undirected type never maps into $\Sigma \bmod \Lambda_2$.

In this section we make these definitions just for the points which are symmetric instances of particles. However, in the next section we mean these definitions for all intersection points, light or dark, symmetric or not.

We define the *fundamental involution* on the space $X_2 = \widehat{X}/\Lambda_2$ to be translation by $(2, P, P)$. This map is the identity mod Λ_1 and an involution mod Λ_2 . Here are the main results of this chapter.

Lemma 9.3 (Vertical) *In each vertical case there exists a sharp container which is a disjoint union of 2 triangles in the $U_1 = U_2$ plane. For the types (P,east) and (Q,west) the vertices of these triangles are*

$$(-2 + P, 1, 1), \quad (-1 + P, 1, 1), \quad (-2 + P, 0, 0)$$

and

$$(P, -1 + P, -1 + P), \quad (-1 + P, -1 + P, -1 + P), \quad (P, P, P)$$

and for the types (P,west) and (Q,east) the triangles are translated by the *fundamental involution*.

Figure 9.1 shows the projection of these sets into the (T, U) plane (for either $U = U_1$ or $U = U_2$) and for the parameter $2/5$. The lightly shaded triangles correspond to (P, west) and (Q, east) and the darkly shaded triangles correspond to (P, east) and (Q, west) . Notice that reflection in the point (P, P, P) swaps these two sets, as does the fundamental involution.

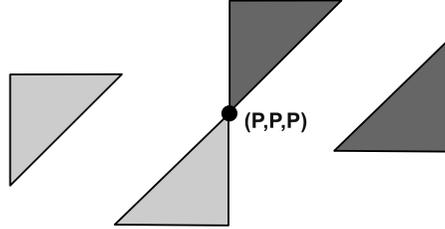


Figure 9.1: The images of vertical symmetric instances

Lemma 9.4 (Horizontal) *In each horizontal case there exists a sharp container which is a disjoint union of two $(1, 1, \sqrt{2})$ right isosceles triangles, one contained in the plane $T = -1$ and one contained in the plane $T = +1$. For the types (P, north) and (Q, south) the vertices of these triangles are*

$$(1, 1, -1 + P), \quad (1, 0, -1 + P), \quad (1, 1, P)$$

and

$$(-1, 1 - P, 1), \quad (-1, 2 - P, 1), \quad (-1, 1 - P, 0)$$

and for the types (P, south) and (Q, north) the new triangles are the images of the ones above under the fundamental involution.

Figure 9.2 shows the projection of the sets corresponding to (P, north) into the (U_1, U_2) plane for the parameter $2/5$. The sets for (Q, south) are the same, and the remaining sets are reflected images of these. The dark triangle lies in the $T = -1$ plane and the light triangle lies in the $T = +1$ plane.

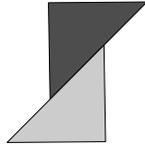


Figure 9.2: The images of horizontal symmetric instances

Remark: The reader can see these sets plotted for any smallish parameter on our computer program. Set the main console to “sanity checks” and the choice to “images by type”.

9.3 Prism Structure

Now we combine the Particle Lemma with the results in the previous section. Every directed half-particle has a symmetric instance, so the results in the previous section apply to every particle. The Particle Lemma combines with the Vertical Lemma and the Horizontal Lemma to prove the following result.

Lemma 9.5 *For every parameter and every type there is a sharp container which is a union of 2 prisms. The vertices of this prism vary linearly with the parameter. The Λ_2 orbit of these prisms consists of an infinite family of pairwise disjoint prisms.*

Proof: A *prism* is a polyhedron which is affinely equivalent to the product of a triangle and an interval. We will give the proof for the particles of type (P,north), but the same argument works in general. The container for the symmetric instances of (P,north) is a union of 2 triangles, one in the plane $T = -1$ and the other in the plane $T = +1$. Call this set Σ_0 . By the Particle Lemma we get a sharp container for all the particles of type (P,north) by considering the union of line segments parallel to $(1, 0, 0)$ and of length $2Q = 4 - 2P$ that are centered on Σ . The set of such points is a union of two prisms. For instance, the prism corresponding to the first triangle in the Horizontal Lemma has vertices

$$(1, 1, -1+P)\pm\eta, \quad (1, 0, -1+P)\pm\eta, \quad (1, 1, P)\pm\eta, \quad \eta = (2-P, 0, 0). \quad (69)$$

We just take the vertices listed in the Triangle Lemma and add $\pm\eta$. The other cases have similar treatments. The crucial fact is that the direction given by the Particle Lemma is always transverse to the triangles given by the Vertical Lemma or the Horizontal Lemma.

In each case, we can see by inspection that the *interiors* of these prisms embed in X_2 . Hence, when we take all the pre-images under the covering map $\widehat{X} \rightarrow X_2$ we get an infinite family of prisms having pairwise disjoint interiors. ♠

Remark: In the vertical case the analysis is slightly misleading. Looking at the Particle Lemma in the vertical case, we see that the picture we get in \mathbf{R}^3 is actually an infinite union of infinite prisms. An infinite prism is the product of a triangle and a line. We get the prisms from the lemma by chopping these infinite prisms by the planes $U_1 = k$ or $U_2 = k$ for odd

integers k . These genuine prisms are not pairwise disjoint but they do have pairwise disjoint interiors.

So far, we have been taking slices one parameter at a time. For each type, we produce a container Σ_P for each parameter $P = 2A/(1 + A)$. Now we take the grand union

$$\Sigma = \bigcup_{P \in [0,1]} \Sigma_P \times \{P\} \subset \mathbf{R}^4. \quad (70)$$

Given the way that the vertices of Σ_P vary with P , the set Σ is a convex integral polytope! The key phenomenon behind this is that the vertices of the prisms vary linearly and the normals to the faces do not change with the parameter.

We find the vertices of the polytope by setting $P = 0$ and $P = 1$ in the formulas. There are no vertices in the intermediate slices because the combinatorics of the prism does not change.

In the example worked out above, the 12 vertices are

$$(1, 1, -1, 0) \pm (2, 0, 0, 0), \quad (1, 0, -1, 0) \pm (2, 0, 0, 0), \quad (1, 1, 0, 0) \pm (2, 0, 0, 0) \\ (1, 1, 0, 1) \pm (1, 0, 0, 0), \quad (1, 0, 0, 1) \pm (1, 0, 0, 0), \quad (1, 1, 1, 1) \pm (1, 0, 0, 0)$$

The grand union of all the triangles in Equation 69 is the convex hull of these 12 vertices.

The affine PET on X_2 gives us a partition of X^2 into 4-dimensional convex integral polytopes. In this section we have constructed a union of 16 convex integral polytopes which determine the nature of the intersection points. The rest of the proof of Theorem 7.5 amounts to lining up these two collections and seeing that they specify the same information.

Notice that the collection of containers produced here is not a partition of X_2 . For instance, there are some point $c \in G$ whose unit integer square consists entirely of dark intersection points. The point $\Phi(c)$ does not lie in any of the containers. Indeed, c has the “all-dark” property if and only if $\Phi(c)$ does not lie in any other containers.

Notice also that the containers can overlap. For instance, there are plenty of grid points c such that $W\Box_c$ contains a light point of type P and a light point of type (P, west) and a light point of type (Q, east) . The point $\Phi(c)$ is contained in the intersection of the corresponding containers.

Nonetheless, the two collections of polytopes will line up exactly in the appropriate sense.

9.4 The Vertical Case

We will use the description of the plaid PET from §7.2. The 4-dimensional version of X_2 is fibered over the (T, P) plane. The (T, P) base space is partitioned into right-angled isosceles triangles which have the property that the fiber over any slice in a given triangle are all combinatorially identical. Call this the *stability property*. Moreover, all such fibers are partitioned into 4×4 grids of rectangles.

Let W_{\leftarrow} denote the union of polytopes which have labels of the form $*W$. These polytopes correspond to tiles which point into the west edge. Likewise, we define W_{\rightarrow} to be the union of polytopes which labels of the form $W*$. These polytopes correspond to tiles which point out of the west edge. The left hand side of figure 9.3 shows these sets over the fiber $(0, 2/5)$.

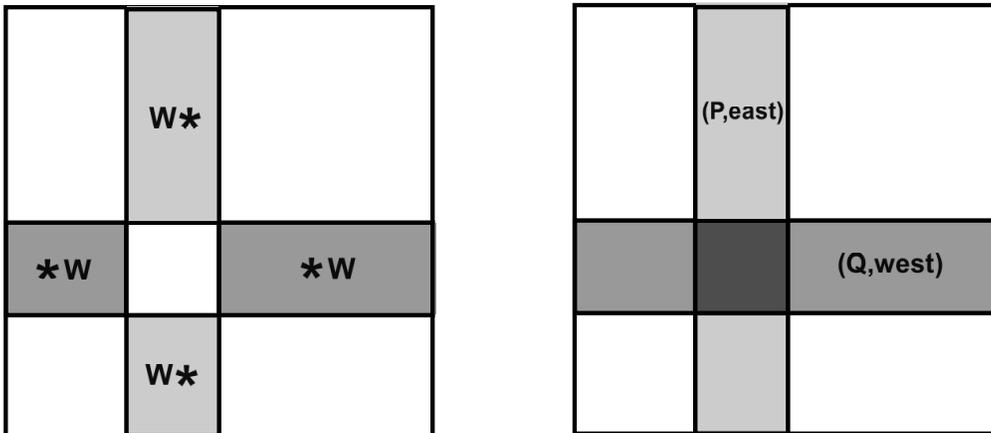


Figure 9.3: The partition compared to the container

Now let's consider the containers. It follows from Statements 1 and 2 of the Particle Lemma that each vertical container intersects each fiber in a rectangle having one of two types.

1. For type P the rectangle has the form $[-1, 1] \times [a, b]$
2. For type Q the rectangle has the form $[a, b] \times [-1, 1]$

By inspection we see that the combinatorics of the intersection does not over the interiors of the triangles in the base partition. We have the same stability. The right hand side of Figure 9.3 shows the intersection of the $(0, 2/5)$ fiber

with the containers. The pattern is the same, except that the intersection of the two strips is precisely the square omitted on the left hand side.

Comparing the two pictures we see that Theorem 7.5 holds in the vertical case for this particular parameter and this particular fiber. Here is a breakdown of what can happen. If $c \in G$ is a point such that $\Phi(c)$ lands in this fiber, then one of 4 things can happen:

- $\Phi(c)$ lands in the white part of the right hand side of Figure 9.3. In this case there are no light points on $W\Box_c$. At the same time the tile assigned to \Box_c by the PET does not involve $W\Box_c$.
- $\Phi(c)$ lands in the central dark square square on the right side of Figure 9.3. In this case there are 2 light points on $W\Box_c$. At the same time the tile assigned to \Box_c by the PET does not involve $W\Box_c$.
- $\Phi(c)$ lands in the container for (P,east) but not in the container for (Q,west). In this case, $W\Box_c$ has one light point and it is directed east. At the same time, $\Phi(c) \in W_{\rightarrow}$, so the PET assigns a tile to \Box_c which points out of c .
- $\Phi(c)$ lands in the container for (Q,west) but not in the container for (P,east). In this case, $W\Box_c$ has one light point and it is directed west. At the same time, $\Phi(c) \in W_{\leftarrow}$, so the PET assigns a tile to \Box_c which points into of c .

In all cases, there is perfect agreement for the two systems of assigning directed connectors in \Box_c .

By direct inspection, we check that such a picture exists in every fiber. This looks like a massive checks, but in all cases we are working with slices of convex integer polytopes, and we have the stability property in all cases. So, we just have to check the agreement in 3 fibers over each triangle in the base partition. This comes to an inspection of 18 pictures like Figure 9.3. We make this check using our program, by direct inspection.

There is another way to think about the inspection. In order to verify the equalities in general, it suffices to verify them for 2 parameters which have the property that a point is mapped into a fiber above each open triangle in the partition. The parameters $2/5$ and $3/8$ would do the job. If there was some mismatch between the two pictures then by Theorem 7.2 there would some $c \in G$ where there was a mismatch in \Box_c between the two schemes. We check directly that there are no such mismatches for these parameters.

9.5 The Horizontal Case

We analyze things in the horizontal case using the same fibration picture as in the vertical case. Figure 9.4 shows the same fiber as in the vertical case, namely the fiber over $(0, 2/5)$. The black region on the right hand side is where the various containers overlap. Inspecting this picture we see once again that both schemes agree on the edge $S\Box_c$ provided that $\Phi(c)$ lands in this fiber.

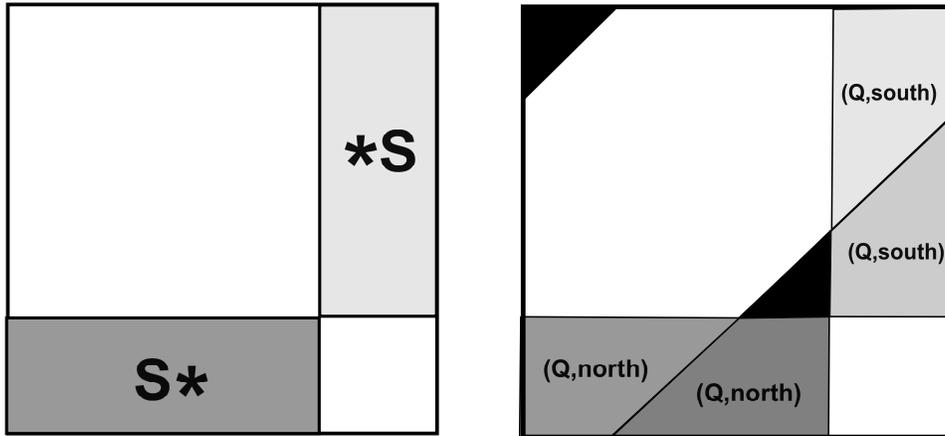


Figure 9.4: The partition compared to the container

There is a second and different way that the various containers can overlap, so we show another picture. Figure 9.5 shows the fiber over $(T, 2/5)$ where $T \approx -1/3$. This point lies in a different isosceles triangle in the base partition. The remaining pictures look like either Figure 9.4 or Figure 9.5.

In Figure 9.5, we use 3 images. The top figure shows the PET picture. This part looks like Figure 9.4. All 4 containers make their appearance in Figure 9.5. The bottom left side of Figure 9.5 shows the containers for $(P,north)$ and $(P,south)$. The bottom right side of Figure 9.5 shows that containers for $(Q,north)$ and $(Q,south)$. Some of the containers from the left overlap with the containers from the right, as is indicated by some crosshatching and shading. Once again, the region covered exactly once is the union of S^* and $*S$, and the directions work out perfectly.

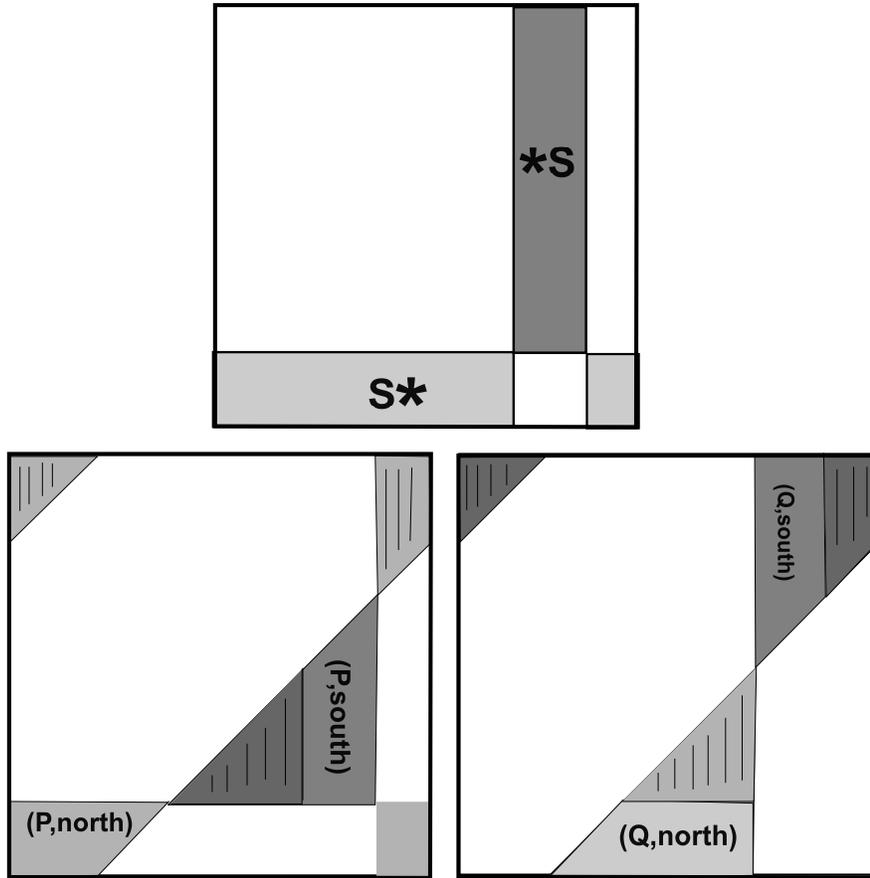


Figure 9.5: The containers in a fiber.

In all cases there is a perfect match-up between the two schemes. As in the vertical case, both schemes are combinatorially stable. Over the same open isosceles triangle in the base partition the combinatorics of the partition does not change. This reduces the check to 18 fibers, as in the vertical case. Again, we make the check by direct inspection.

This completes the proof in the horizontal case.

10 Proof of the Vertical Lemma

10.1 Outline

The goal of this chapter is to prove The Vertical Lemma. We will do it in 5 steps.

1. We will use symmetry to cut down the amount of work we have to do. By symmetry we will only have to look at the case (P,east).
2. We will translate the picture so that we are trying to prove something more symmetric. The Vertical Lemma associates a point at the center of the unit integer square with a point on the east edge of the square. We get a more symmetric picture when we work with the centers of the east edges rather than the centers of the squares.
3. If we work mod Λ_1 rather than mod Λ_2 then the two cases (P,east) and (P,west) get amalgamated into one case, and the corresponding statement is simpler to prove. We will first formulate and prove our result mod Λ_1 .
4. We present some alternative formulas for the map $\Phi : G \rightarrow X_1$ and $\Phi : G \rightarrow X_2$ which will help with our calculations.
5. We will make a careful study of the congruences and show that the Λ_1 -based result implies the Λ_2 -based result.

10.2 Using Symmetry

Let \square_c denote the unit integer square centered at a grid point c . Let $W\square_c$ and $S\square_c$ denote the east and south edges of \square_c .

Lemma 10.1 *The same sharp container works for the types (P,east) and (Q,west) simultaneously and likewise the same sharp container works for (P,west) and (Q,east) simultaneously.*

Proof: Suppose c lies on a horizontal midline H . Reflection in H interchanges the two light points on $W\square_c$, referses their types, and reverses their directions. Hence the two types (P,east) and (Q,west) always appear together in this situation, and likewise for the types (P,west) and (Q,east). ♠

Lemma 10.2 *If a given sharp contained works for (P, east) then the image of this container under the fundamental involution works for (P, west) .*

Proof: The map $(x, y) \rightarrow (x, y) + (0, \omega)$ preserves the types of the particles and reverses their directions. Also, we have

$$\Phi(x, y + (0, \omega)) = \Phi(x, y) + (2, P - 2, P - 2) \bmod \Lambda_2. \quad (71)$$

This proves the last statement of this lemma. ♠

In view of these symmetry results, it suffices to prove the Vertical Lemma for the case (P, east) .

10.3 Translating the Picture

We modify the original construction. Keeping the same notation as above, we now associate to $z \in W\square_c$ the point $\Phi(c')$ where c' is the center of $W\square_c$. What is more natural about this construction is that z and c' lie on the same unit integer segment.

We call a subset $\Sigma' \subset X_2$ a *translated sharp container* for (P, east) if it has the following property: $\Phi(c') \in \Sigma'$ if and only if one of the two intersection points on the unit segment centered at c' has type (P, east) . Looking at Equation 45, we see that Σ is a sharp container for (P, east) if and only if

$$\Sigma' = \Sigma - (P, P, P)$$

is a translated sharp container for (P, east) . The translated sharp container is easier to work with.

Here is an equivalent formulation of the Vertical Lemma.

Lemma 10.3 *In each vertical case there exists a modified sharp container for (P, east) which is a union of the 2 triangles with vertices*

$$(-2, 1 - P, 1 - P), \quad (-1, 1 - P, 1 -), \quad (-2, -P, -P)$$

and

$$(0, -1, -1), \quad (-1, -1, -1), \quad (0, 0, 0)$$

Lemma 10.3 is just like the Vertical Lemma except that we have

10.4 Forgetting the Directions

The left half of 10.1 shows a repeat of Figure 9.1 except with the point (P, P, P) translated to $(0, 0, 0)$. This is our translated sharp container. We are drawing this picture in the (T, U_j) plane. The left half of the picture is a subset of $[-2, 2] \times [-1, 1]$.

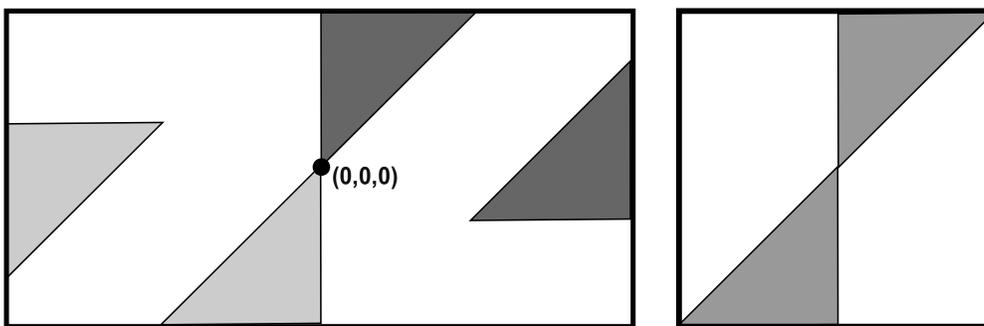


Figure 10.1: The images of vertical symmetric instances

The right half of the picture shows the image of the left half of the picture under the quotient map $X_2 \rightarrow X_1$. This quotient map identifies the light and dark shaded triangles using the fundamental involution, which is translation by $(2, P, P)$. This is the set of points $(T, U_1, U_2) \subset [-1, 1]^3$ such that

- $U_1 = U_2$
- U_j and T have the same sign.
- $|T| \leq |U_j|$.

We will prove that the union of two triangles on the right side of the Figure is a translated sharp container for light points of type P when we work mod Λ_1 . That is, if c is the center of a vertical unit integer segment then $\Phi(c)$ lies in this set if and only if one of the light points on this edge has type P.

After we deal with the Λ_1 case, we will again work mod Λ_2 and show that the restriction of Θ to the subset of the edge centers corresponding to (P, east) has its image in the left half of the range, $[-2, 0] \times [-1, 1]^2$ and the restriction to the edge centers corresponding to (P, west) has its image in the right half. This is how the Λ_1 result implies the Λ_2 result. For this latter purpose, we just have to look at the first coordinate of Φ mod Λ_2 .

10.5 Some Useful Formulas

Let $[x]_{2n}$ denote the representative of $x \bmod 2n\mathbf{Z}$ in $[-n, n)$. First we derive some alternate formulas for the map Φ .

Lemma 10.4 *Let $c = (x, y)$. Then $\Phi(c) = (T, U_1, U_2) \bmod \Lambda_2$, where*

- $T(c) = [2Px + 2y]_4$.
- $b(c) = \frac{1}{2}PT(x, y)$.
- $U_1(c) = [PQx + b(x, y) - Py]_2$.
- $U_2(c) = [PQx + b(x, y) + Py]_2$.

Proof: Let Φ' be the map above. $T(c)$ is just the expression for the first coordinate of $\Phi(c)$. Hence, $\Phi = \Phi'$ in the first coordinate.

Φ' is locally affine on $(x, y) \in \mathbf{R}^2$ as long as $2Px + 2y \notin 4\mathbf{Z} + 2$. Suppose we evaluate Φ' on two points of (x, y_1) and (x, y_2) , with $y_2 = y_1 + \epsilon$ for some small ϵ . If $[2Px + 2y_1]_2 \approx 2$ and $[2Px + 2y_2]_2 \approx -2$ then $b(x, y_1) \approx P$ and $b(x, y_2) \approx -P$. But then

$$\Phi'(x, y_2) - \Phi'(x, y_1) \approx (4, 2P, 2P) \in \Lambda_2.$$

Letting $\epsilon \rightarrow 0$ we see that Φ' is continuous mod Λ_2 .

Since Φ' is continuous on \mathbf{R}^2 , locally affine away from a countable set of lines, and has constant linear part, Φ' is locally affine on all of \mathbf{R}^2 . But so is Φ . Using the identity $PQx + P^2x = 2Px$, we check that $\Phi = \Phi'$ in a neighborhood of the origin. Since $\Phi = \Phi'$ agree near the origin, these two locally affine maps agree everywhere. ♠

If y is a half integer, and we work mod Λ_1 , then we get the same formula, except that we can use the simpler expression

$$T(x, y) = [2Px + 1]_2. \tag{72}$$

in place of $[2Px + 2y]_4$.

Now we introduce some functions which will help us decide whether an intersection point is a light point. Let \mathbf{Z}_0 and \mathbf{Z}_1 denote the sets of even and odd integers respectively. We define the following 4 functions:

- $F_H(x, y) = [2Py]_2$
- $F_V(x, y) = [2Px]_2$
- $F_P(x, y) = [Py + P^2x + 1]_2$
- $F_Q(x, y) = [Py + PQx + 1]_2$

We shall never be interested in the inverse images of \mathbf{Z}_1 , so there is never any ambiguity in the definition.

Lemma 10.5 *A vertical intersection point z lying on a slanting line of type P is a light point if and only if $F_P(z)$ and $F_V(z)$ have the same sign and $|F_P(z)| < F_V(z)$. The same result holds with Q replacing P .*

Proof: The value $\omega F_V(z)$ computes the signed capacity of the vertical line containing z . The function F_P is constant on lines of slope $-P$ and $\omega F_P(z)$ computes the signed mass of the slanting line of slope $-P$ through V . ♠

Lemma 10.6 *A horizontal intersection point z lying on a slanting line of type P is a light point if and only if $F_P(z)$ and $F_H(z)$ have the same sign and $|F_P(z)| < F_H(z)$. The same result holds with Q replacing P .*

Proof: This has the same kind of proof. ♠

10.6 The Undirected Result

In this section we prove the Λ_1 version of Lemma 10.3. Let $c = (x, \omega/2)$ be the center of some vertical unit integer segment centered on the horizontal midline of a block. Let $z = (x, y)$ denote be the intersection of a slanting line of type P with the vertical unit integer segment centered at c .

Lemma 10.7 $y = \omega/2 + T(c)/2$.

Proof: We have

$$y \in (\omega/2 - 1/2, \omega/2 + 1/2). \quad (73)$$

The point z lies on a line of slope $-P$ which has an integer y -intercept. Hence $y + Px \in \mathbf{Z}$. By construction, $Px + y$ is the integer nearest $Px + \omega/2$. Hence

$$y + Px = Px + \omega/2 - [Px + \omega/2]_1.$$

Subtracting off Px , we get

$$y = \omega/2 - [Px + \omega/2]_1. \quad (74)$$

Since $[t]_1 = [2t]_2/2$, we get

$$y = \omega/2 - \frac{1}{2}[2Px + \omega]_2 =^* \omega/2 - \frac{1}{2}[2Px + 1]_2 = \omega/2 - T(c)/2. \quad (75)$$

The starred equality uses the fact that $\omega - 1 \in 2\mathbf{Z}$. ♠

Now we are going to do a trick. Rather than work with the point z we will work with the point

$$z' = (x, \omega/2 + T(c)/2). \quad (76)$$

Reflection in the horizontal midline of the block swaps z and z' . By symmetry, z is a light point of type P if and only if z' is a light point of type Q.

Now we use Lemma 10.5. We have

$$F_V(z') = [2Px]_2 = [T(c) + 1]_2. \quad (77)$$

Lemma 10.8

$$F_Q(z') = [U_1(c) + 1]_2. \quad (78)$$

Proof: Using Lemma 10.4 mod Λ_1 , and remembering that $c = (x, \omega/2)$,

$$\begin{aligned} F_Q(z') &= [Py + PQx + 1]_2 = \\ &= [P\omega/2 + PT(c)/2 + PQx + 1]_2 = \\ &= [PQx + b(c) + P\omega/2 + 1]_2 = \\ &= [U_2(c) + 1]_2 = [U_1(c) + 1]_2. \end{aligned} \quad (79)$$

The last equality comes from the fact that we already know $U_1(c) = U_2(c)$. ♠

From Equations 77 and 78 we see that $F_V(z')$ and $F_Q(z')$ have the same sign iff $T(c)$ and $U_1(c)$ have the same sign and $|F_Q(z')| < |F_V(z')|$ iff $|T(c)| < |U(c)|$. Hence z' is a light point of type P if and only if $\Phi(c)$ lies in the container described above.

10.7 Determining the Directions

Now we deduce the directed result from the undirected result. We keep the same notation but work mod Λ_2 . The first coordinate of $\Phi(c)$ is

$$[2Px + \omega]_4. \quad (80)$$

We want to show that this quantity is positive iff z has type (P,east).

The vertical line containing c and z either has positive sign or negative sign. We first consider the case when this line has positive sign. This is to say that $[4px]_{2\omega} > 0$. Dividing through by ω , we get

$$[2Px]_2 > 0. \quad (81)$$

Let y_0 denote the y -intercept of the line of slope $-P$ through z . By From the argument in Lemma 1.3, the point z has type (P,east) iff y_0 is even. We have

$$y_0 = Px + y.$$

y_0 is the integer nearest $Px + \omega/2$. So, the integer nearest $Px + \omega/2$ is even. Hence the even integer nearest $2Px + \omega$ is congruent to 0 mod 4. In summary,

1. $[2Px]_2 > 0$
2. The even integer nearest $2Px + \omega$ is congruent to 0 mod 4.

The left half of Figure 10.2 explains why these two conditions imply that $[2Px + \omega]_4 > 0$. The thick lines above the interval show the constraints given by our first condition and the thick lines below show the constraints given by the second.



Figure 10.2: How the constraints force the sign.

When our vertical line has negative sign, the argument is similar. The right side of Figure 10.2 shows this case.

11 Proof of The Horizontal Lemma

11.1 Using Symmetry

In this chapter we prove the Horizontal Lemma. We follow the same pattern of proof as for the Vertical Lemma. We begin by reducing the number of cases we have to study.

Lemma 11.1 *The same sharp container works for the types (P, south) and (Q, north) simultaneously and likewise the same sharp container works for (P, north) and (Q, south) simultaneously.*

Proof: As discussed in §1.4 (the Midline Case), the edge $S\Box_c$ contains two intersection points at its midpoint μ when c is contained in a vertical midline. One could say that both these points have both types P and Q . What is important here are the directions.

Let L_P and L_Q respectively denote the slanting lines of slope $-P$ and $-Q$ through μ . These lines have the same sign because μ is a light point with respect to both. In 1.3 we proved that L'_Q and L_Q give the same direction to μ . Also, from the second definition of the plaid model, L_P , L_Q , and L'_Q all have the same sign. Since the first coordinate of μ has the form $k\omega + \omega/2$ and since $P + Q = 2$, we see that the difference in the y -intercepts of L_P and L'_Q is an odd integer. Hence L_P and L'_Q have opposite parity. Hence these lines give the opposite direction to μ . So, for these symmetric instances, the types (P, north) and (Q, south) always appear together and the type (P, south) and (Q, north) always appear together. ♠

In the horizontal case, we will use symmetry in a slightly different way. Let σ denote our (putatively) sharp container for (P, north) and let ν denote our (putatively) sharp container for (P, south) . We have $\sigma = \sigma_+ \cup \sigma_-$ where σ_\pm is the half of σ contained in the plane $T = \pm 1$. We have the same decomposition for ν .

Lemma 11.2 *Suppose σ_+ and ν_+ are respectively sharp containers for those light points of type (P, south) and (P, north) whose characters lie in the plane $T = 1$. Then σ is a sharp container for (P, south) and ν is a sharp container for (P, north) .*

Proof: We just have to show that σ_- and ν_- are sharp containers for those light points of type (P,south) and (P,north) whose characters lie in the plane $T = -1$. One can check from the definitions of our sets in the Horizontal Lemma that

$$\sigma_- = \nu_+ - (2, P, P), \quad \nu_- = \sigma_+ - (2, P, P), \quad \text{mod } \Lambda_2. \quad (82)$$

Let z be such a point. We will treat the case when z has type (P,south). The other case is similar.

Let $z' = z + (0, \omega)$. Let c and c' denote the grid points associated to z and z' respectively. We have $c' = c + (0, \omega)$. Translation by $(0, \omega)$ preserves the (undirected) types and reverses the directions. Hence z' has type (P,north). By assumption, $\Phi(c') \in \nu_+$. But

$$\Phi(c) = \Phi(c') - (2, P, P) \text{ mod } \Lambda_2.$$

Hence $\Phi(c) \in \nu_+ - (2, P, P)$. Our claim now follows from Equation 82. ♠

11.2 Translating the Picture

We modify the construction as in the vertical case. This time we work with the centers of horizontal unit integer segments. When we replace the center c of a unit integer square by the center $c' = c - (0, 1/2)$ of the south edge of that square, we translate the sets σ_+ and ν_+ by $(-1, 0, -P) \text{ mod } \Lambda$. The translated set σ'_+ is the triangle with vertices

$$(0, 1, -1), \quad (0, 0, -1), \quad (0, 1, 0) \quad (83)$$

To get the set μ'_+ we add $(2, P, P)$ to the second triangle in the Horizontal Lemma, then subtract $(1, 0, P)$, then take the result mod $2\mathbf{Z}$. The answer is as nice as possible:

$$(0, -1, 1), \quad (0, 0, 1), \quad (0, -1, 0) \quad (84)$$

Figure 11.1 shows these two triangles. The lightly shaded triangle is σ_+ and the darkly shaded triangle is σ_- . We are drawing the picture in the $T = 0$ slice. The picture lies in $[-1, 1]^2$.

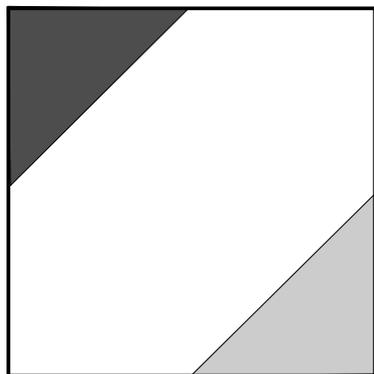


Figure 11.1: How the constraints force the sign.

11.3 A Technical Lemma

We call a pair of points $(a, b) \in \mathbf{R}^2$ a *good pair* if $[a]_2 \leq [b]_2$ and $[a_2]$ has the same sign as $[b_2]$. The set of good pairs in $[-1, 1]$ is exactly the set shown in the right hand side of Figure 10.1.

Lemma 11.3 *Let $b_1 = a_1 - a_2$ and $b_2 = a_1 - 1$. Then (a_1, a_2) is a good pair if and only if (b_1, b_2) is a good pair.*

Proof: Since the affine transformation $T(x, y) = (x - y, x - 1)$ preserves \mathbf{Z}^2 , the pair (a_1, a_2) satisfies the non-integrality condition if and only if the pair (b_1, b_2) does. If this lemma is true for the inputs (a_1, a_2) it is also true for the inputs $(a_1 + 2k_1, a_2 + 2k_2)$ for any integers k_1, k_2 . For this reason, it suffices to consider the case when $a_1, a_2 \in (-1, 1)$. From here, an easy case-by-case analysis finishes the proof. For instance, if $0 < a_1 < a_2$ then $0 > b_1 > b_2$. The other cases are similar. ♠

11.4 The Undirected Result

As in the vertical case, we first work mod Λ_1 . That is, we forget about the directions. We show that the union of the two triangles in the previous section is a translated sharp container mod Λ_1 for the symmetric instances of light particles. Again, all such light points on slanting lines of type P and of type Q. So, when we ignore the directions, there are no distinctions to be made for an intersection point, except whether it is light or dark.

Now we observe a feature of the horizontal case. For the symmetric instances of particles, the intersection points already occur at the midpoints of the horizontal edges. This is to say that the light point z and the center point c coincide. We call this common point $c = (x, y)$. By assumption

$$x = k\omega + \omega/2, \quad y \in \mathbf{Z}. \quad (85)$$

Working mod Λ_1 , we know that the first coordinate of $\Phi(c)$ is 0. By Lemma 10.4 the other two coordinates are

$$U_1(c) = [PQx - Py]_2, \quad U_2(c) = [PQx + Py]_2. \quad (86)$$

Referring to the useful formulas given in the previous chapter, we have

$$F_P(c) = F_Q(c) = [PQx + Py + 1]_2, \quad F_H(c) = [2Py]_2. \quad (87)$$

Setting $U_1 = U_1(c)$, etc., we see that mod $2\mathbf{Z}$ we have

$$U_1 = F_P - F_H + 1, \quad U_2 = F_P + 1. \quad (88)$$

For the moment let $U'_j = U_j + 1 \pmod{2\mathbf{Z}}$. The point c is a light point if and only if $|F_P| < |F_H|$ and these two quantities have the same sign. By Lemma 11.3, the set of points (U'_1, U'_2) corresponding to the light points is the same set: It is the set where $|U'_1| < |U'_2|$ and the two quantities have the same sign. But the set in Figure 11.1 is the translate of this set by the vector $(1, 1) \pmod{2\mathbf{Z}}$. This proves the undirected case of the horizontal lemma.

11.5 Determining the Directions

Now we go back to working mod Λ_2 . Suppose that $c = (x, y)$ is a light point of type (P,south). It suffices to show that the condition

$$[2Px + 2y]_4 = 0 \quad (89)$$

implies that

$$[2Py]_2 = [U_2 - U_1]_2 \leq 0. \quad (90)$$

In other words, we add the coordinates, take the representative mod $2\mathbf{Z}$ in $[-1, 1)$ and check that it is non-negative. We don't have to worry about the boundary case because from the analysis in §7.3 the point (U_1, U_2) does not lie in the boundaries of our triangles.

The sign of the horizontal line through c is given by $[2Py]_2$. The direction of the slanting line of slope $-P$ through c is determined by the sign of $[2Py]_2$ and the parity of the y -intercept of the line of slope $-P$ through c . This is the point

$$y_0 = y + Px. \tag{91}$$

Equation 89 says that $2y_0$ is divisible by 4. Hence y_0 is even. But then the direction of c determines the sign of $[U_2 - U_1]_2$. When c has type (P,north), the expression is non-negative as in Equation 90.

This completes the proof of the Horizontal Lemma.

Part III

The Graph Master Picture Theorem

12 The Arithmetic Graph

This begins the Part 3 of the monograph. In this part, we describe the arithmetic graph for outer billiards on kites and the Graph Master Picture Theorem.

12.1 Special Orbits and the First Return Map

We will describe the situation when $p/q \in (0, 1)$ is rational. We need not take pq even. Let $A = p/q$. We consider outer billiards on the kite K_A , which has vertices $(0, 1)$, $(0, -1)$, $(-1, 0)$, and $(A, 0)$. Recall from the introduction that the *special orbits* are the orbits which lie on the set S of horizontal lines having odd integer y -intercepts.

Let ψ_A denote the second iterate of the outer billiards map. This map is a piecewise translation. Let Ψ_A denote the first return map of ψ_A to the union

$$\implies = \mathbf{R}_+ \times \{-1, 1\} \subset S \quad (92)$$

Lemma 12.1 *Every special orbit is combinatorially identical to the orbit of a point of the form*

$$\left(2mA + 2n + \frac{1}{q}, \pm 1\right). \quad (93)$$

Here $m, n \in \mathbf{Z}$ and the left hand side of the equation is meant to be positive.

Proof: This is proved in [S1]. Here is a sketch of the proof. What is going on is that ψ_A permutes the intervals of the form

$$S = \bigcup_{\ell, m, n \in \mathbf{Z}} \left(\frac{2m}{q} + n, 2\ell + 1\right)$$

and so every orbit is combinatorially identical to the orbit of a center point of one of these intervals. Moreover, all such orbits intersect the set \implies . When the orbit of a center point of an interval in the partition intersects \implies , it does so in a point of the form given in Equation 93. ♠

Remark: The same result holds if $1/q$ in Equation 93 is replaced by any number in $(0, 2/q)$. The choice of $1/q$ seems the most canonical.

12.2 The Arithmetic Graph

Suppose that m_0, n_0 are integers such that $2m_0A + 2n_0 \geq 0$. Then, by definition, there are integers m_1, n_1 , with $2m_1A + 2n_1 \geq 0$ such that

$$\Psi_A\left(2m_0A + 2n_0 + \frac{1}{q}, 1\right) = \left(2m_1A + 2n_1 + \frac{1}{q}, \epsilon\right), \quad (94)$$

where $\epsilon \in \{-1, 1\}$ is given by

$$\epsilon = (-1)^{m_0+m_1+n_0+n_1}. \quad (95)$$

Reflection in the x -axis conjugates Ψ_A to Ψ_A^{-1} , so we also have

$$\Psi_A^{-1}\left(2m_0A + 2n_0 + \frac{1}{q}, -1\right) = \left(2m_1A + 2n_1 + \frac{1}{q}, -\epsilon\right), \quad (96)$$

Actually, we won't end up caring about ϵ .

Figure 12.1 shows a cartoon of what Equation 94 looks like geometrically when $\epsilon = -1$.

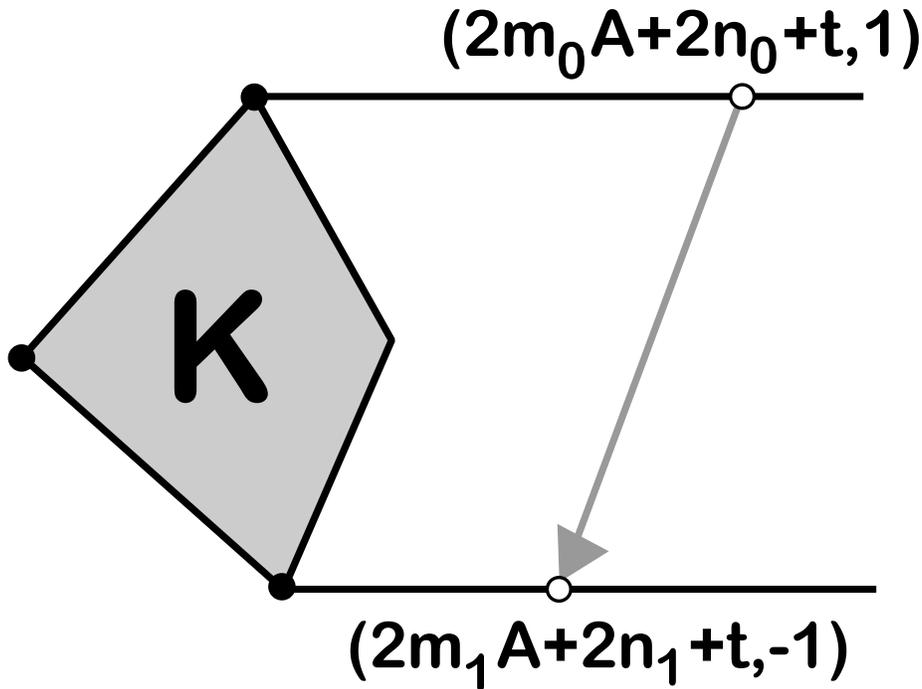


Figure 12.1: The arithmetic graph construction.

We form a graph Γ_A whose vertices are \mathbf{Z}^2 by joining (m_0, n_0) to (m_1, n_1) by an edge if and only if these points are related as in Equation 94. We proved in [S1] that all the edges of Γ_A have length at most $\sqrt{2}$. That is, there are just 8 kinds of edges. We also proved that Γ_A is a union of pairwise disjoint embedded polygonal paths. All these paths are closed when pq is even. Our proof of the Quasi-Isomorphism will reprove these statements about the arithmetic graph.

The nontrivial components of Γ all lie in the half plane above the line L of slope $-A$ through the origin. The map

$$f(m, n) = \left(2Am + 2n + \frac{1}{q}, (-1)^{m+n} \right) \quad (97)$$

carries each component of Γ_A to a different special orbit. The image of this map is “half” the special special orbits, in the sense that every special orbit, or its mirror, is represented. The *mirror* of a special orbit is its reflection in the x -axis. If we wanted to get the other half of the special orbits, we would use the map $\rho \circ f$, where ρ is reflection in the x -axis.

It is possible to extend Γ_A in a canonical way so that it fills the entire plane, and not just the half plane. We want to do this so that the Quasi-Isomorphism Theorem is true as stated. There are two ways to do this, and they give the same answer.

1. Dynamically, we can consider the first return map to the negative-pointing rays and make the same construction.
2. Using the classifying space picture described below, we can simply take the domain to be all of \mathbf{Z}^2 rather than just the portion of \mathbf{Z}^2 above L .

We will take the second approach below. We call this extended version of Γ_A the *arithmetic graph* at the parameter A .

Remarks:

- (i) As with the plaid model, there is a question as to whether or not we orient the components of the arithmetic graph. The orientation comes naturally from the outer billiards dynamics. We generally will not pay attention to this orientation.
- (ii) In §15.7 we will give a definition of the arithmetic graph which works for any polygon without parallel sides. The construction given above is a special case of the general construction.

12.3 The Canonical Affine Transformation

The Quasi-Isomorphism Theorem compares the plaid model with a certain affine image of the arithmetic graph. In this section we describe the affine map that we need to apply to the arithmetic graph to make the direct comparison.

Recall that $A = p/q$ and $\omega = p + q$. The affine map from the Quasi-Isomorphism Theorem is given by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{A+1} \begin{pmatrix} A^2 + A & A+1 \\ -A^2 + 2A + 1 & -2A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \frac{1}{2q} \\ -\frac{-1}{2q} + \frac{1}{p+q} + \tau \end{pmatrix}. \quad (98)$$

Here τ is the solution in $(0, \omega)$ to the equation $2p\tau \equiv 1 \pmod{\omega}$. The linear part of the map T is defined for irrational parameters as well as rational parameters, but the map itself is only defined when p/q is an even rational.

Normalized Arithmetic Graph: The canonical affine transformation T is the implied affine map in the Quasi-Isomorphism Theorem. We define the *normalized arithmetic graph* to be the T -image of the arithmetic graph. The Quasi-Isomorphism Theorem says that the plaid model and the normalized arithmetic graph are 2-quasi-isometric at each parameter.

The inverse of the canonical affine transformation is

$$T^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{(A+1)^2} \begin{pmatrix} 2A & A+1 \\ -A^2 + 2A + 1 & -A^2 - A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2p+2q} \begin{pmatrix} -1 - 2q\tau \\ -1 + 2p\tau \end{pmatrix}. \quad (99)$$

Again, the linear part is defined even for irrational parameters. However, the entire map is only defined for even rationals.

The Graph Grid: We define the *graph grid* to be the grid $T(\mathbf{Z}^2)$. The vertices of the normalized arithmetic graph lie in the graph grid. A calculation shows that $\det(dT) = 1 + A$, so the graph grid has co-area $1 + A$, as mentioned in the introduction. What makes T canonical is that $T(\mathbf{Z}^2)$ has rotational symmetry about the origin.

The Anchor Point: Let

$$\zeta = \left(\frac{1+A}{2}, \frac{1-A}{2} \right). \quad (100)$$

We compute

$$T^{-1}(\zeta) = \left(\frac{1}{2}, \frac{1}{2} \right) + \left(\frac{-1-2q\tau}{2\omega}, \frac{-1+2p\tau}{2\omega} \right).$$

The second summand is a vector having half-integer coordinates. Hence, $T^{-1}(\zeta) \in \mathbf{Z}^2$. Hence $\zeta \in T(\mathbf{Z}^2)$. We call ζ the *anchor point*. Given the existence of ζ , we can redefine the graph grid to be the translate of $dT(\mathbf{Z}^2)$ which contains ζ . This allows us to define the graph grid even at irrational parameters, even though the canonical affine transformation is only defined for even rational parameters. As A varies from 0 to 1, the grid $T(\mathbf{Z}^2)$ interpolates between the grid of half integers and the grid of integers whose coordinates have odd sum, and each individual point travels along a hyperbola or straight line.

Distinguished Edges: Say that a *distinguished edge* in the grid graph is one connecting distinct points of the form

$$T(\zeta), \quad T(\zeta + (i, j)), \quad i, j \in \{-1, 0, 1\}. \quad (101)$$

Let $\mathcal{F}(i, j)$ denote the family of distinguished edges corresponding to the pair (i, j) . In [S0] we proved that the arithmetic graph is embedded, and each edge is one of the 8 shortest vectors in \mathbf{Z}^2 . Thus, the edges of the normalized arithmetic graph are all distinguished.

Distinguished Lines: Say that a *distinguished line* is a line that contains a distinguished edge of the grid graph. If these lines are parallel to edges in $\mathcal{F}(i, j)$ we say that the lines have *type* (i, j) .

12.4 Geometry of the Graph Grid

In this section we prove a number of statements about the geometry of the graph grid. The word *square* always means integer unit square, as in the plaid model.

Lemma 12.2 (Grid Geometry) *The following is true at each parameter.*

1. No point of the grid graph lies on the boundary of a square.
2. Two points of the grid graph cannot lie in the same square.
3. A union of 3 horizontally consecutive squares intersects the grid graph.
4. A union of 2 vertically consecutive squares intersects the grid graph.
5. Two graph grid points in consecutive squares are always connected by a distinguished edge.
6. If two parallel distinguished lines intersect the interior of the same edge e of a square, then the lines have type $(-1, 1)$ and e is horizontal.
7. The slopes of the distinguished edges are never in $\{-1, 1, 0, \infty\}$.

Proof of Statement 1: We just need to prove, for each $(x, y) \in \mathbf{Z}^2$, that neither coordinate of $T(x, y)$ is an integer. The two coordinates of $T(x, y)$ are

$$\frac{1}{2q} + \frac{px}{q} + y, \quad \tau - \frac{q-p}{2q(p+q)} + \frac{-2p^2x + 4pqx + 2q^2x - 4pqy}{2q(p+q)}$$

The first number is always $k/(2q)$ where K is an odd integer. The second number is always $n/(2q(p+q))$, where n is odd. ♠

Lemma 12.3 Suppose v is a vector with $\|v\| < \sqrt{2}$. Then $\|dT^{-1}(v)\| < 2$.

Proof: For any matrix M , we have $\|M(v)\| \leq \|M\| \|v\|$. Here $\|M\|_2$ is the L_2 norm of M . Setting $M = dT^{-1}$ we get

$$\|M\|_2^2 = \sum_{ij} M_{ij}^2 = \frac{2(1 + 3A + 4A^2 - A^3 + A^4)}{(1 + A)^4} \leq 2$$

The inequality on the right, which holds for all $A \in [0, 1]$, is an exercise in calculus: We check that the derivative of this expression does not vanish in $[0, 1]$, and then we evaluate at $A = 0$ and $A = 1$ to get the bound on the right. The final result is that $\|M\|_2 \leq \sqrt{2}$. ♠

Proof of Statement 2: Call a vector in $dT(\mathbf{Z}^2)$ *bad* if both its coordinates are less than 1 in absolute value. Thanks to the previous result, if this lemma is false then there is some bad vector in $dT(\mathbf{Z}^2)$. A bad vector has length less than $\sqrt{2}$. Hence, by the previous result, a bad vector must have the form $dT(\zeta)$, where ζ is one of the 8 shortest nonzero vectors in \mathbf{Z}^2 . By symmetry we just have to check 4 out of the 8, namely:

- $dT(1, 0) = \left(A, \frac{1+2A-A^2}{1+A} \right)$.
- $dT(0, 1) = \left(1, \frac{-2A}{1+A} \right)$.
- $dT(1, 1) = (1 + A, 1 - A)$.
- $dT(1, -1) = \left(-1 + A, \frac{1+4A-A^2}{1+A} \right)$.

Here we have set $A = p/q \in (0, 1)$. In all cases, we can see that least one coordinate is at least 1 in absolute value. ♠

Proof of Statement 3: It suffices to prove that every translate of $T^{-1}(R)$ intersects \mathbf{Z}^2 . The vertices of $T^{-1}(R)$ are

$$R_0 = (0, 0), \quad R_1 = \left(1, -\frac{2A}{1+A} \right), \quad R_2 = \left(3A, \frac{3+6A-3A^2}{1+A} \right), \quad R_3 = R_1 + R_2.$$

Clearly R_1 lies below the line $y = 0$. A bit of calculus shows that R_2 and R_3 both lie above the line $y = 2$. Furthermore, any line of the form $y = h$ with $h \in [0, 2]$ intersects $T^{-1}(R)$ in a segment of width

$$\frac{1 + 2A + A^2}{1 + 2A - A^2} > 1.$$

This is sufficient to see that every translate of $T^{-1}(\mathbf{R})$ intersects \mathbf{Z}^2 . ♠

Proof of Statement 4: Let R be a union of two vertically consecutive squares. Let C_R denote the infinite column of integer unit squares containing R . Note that $T(0, 1) = (1, -P)$. Say that a *distinguished array* is an infinite set of points of the form $T(m_0, n)$, where m_0 is held fixed and $n \in \mathbf{Z}$. Because the first coordinate of $T(0, 1)$ is 1, we see that every row intersects C_R in one point.

Let ρ_m denote the m th row of points. Let y_m denote the y -coordinate of the point v_m where ρ_m intersects C_R . Looking at the formulas above, observe that the difference in the x coordinates of $dT(1, 0)$ and $dT(1, -1)$ is 1 and that both y coordinates are in $[1, 2]$. For this reason, one of the two points $v_m + dT(1, 0)$ or $v_m + dT(1, -1)$ lies on a row above v_m and moreover the vertical distance between these points is less than 2. Hence $|h_{m+1} - h_m| < 2$ for all m . But then R contains some v_m . ♠

Proof of Statement 5: Let

$$f = v_1 - v_2 = (f_1, f_2) = dT(i, j).$$

When the squares are stacked on top of each other, we have the constraints $|f_1| < 1$ and $|f_2| = 2$. When the squares are stacked on top of each other, we have the constraints $|f_1| < 2$ and $|f_2| = 1$. Both cases give $\|f\| < \sqrt{5}$. This combines with $\|dT^{-1}\|_2 < \sqrt{2}$ to show that $i^2 + j^2 < 10$. An explicit case-by-case rules everything out but

- $(i, j) \in \{\pm(1, 0), \pm(1, -1)\}$ in the first case.
- $(i, j) \in \{\pm(0, 1), \pm(1, 1)\}$ in second case.

These cases do actually occur. ♠

Proof of Statement 6: Let us first consider in detail the case when L_1 and L_2 have type $(1, 0)$ and the edge e is vertical. To rule out this case, we just need to intersect two adjacent lines with the y -axis and see that the distance between the intersection points is at least 1 unit. Let dT be the linear part of T . Two consecutive lines are given by

$$(1 - s)dT(0, 0) + sdT(1, 0), \quad (1 - s)dT(0, 1) + sdT(1, 1).$$

The first line contains $(0, 0)$. To see where the second line intersects the y -axis, we set the first coordinate equal to 0 and solve for t . This gives $t = -1/A$. Plugging this into the equation, we see that the y intercept is $-1 - 1/A$. Hence, the vertical distance is $1 + 1/A$, a quantity that always exceeds 1. We record this information by writing $d(1, 0, V) = 1 + 1/A$. Following the same method, we do the other 7 cases. Here is the result:

$$d(1, 0, V) = 1 + \frac{1}{A}, \quad d(1, 0, H) = \frac{1 + 2A + A^2}{1 + 2A - A^2}.$$

$$\begin{aligned}
d(0, 1, V) &= 1 + A, & d(0, 1, H) &= \frac{1 + 2A + A^2}{2A}. \\
d(1, 1, V) &= 1, & d(1, 1, H) &= \frac{1 + A}{1 - A}. \\
d(-1, 1, V) &= \frac{1 + A}{1 - A}, & d(-1, 1, H) &= \frac{1 + 2A + A^2}{1 + 4A - A^2}.
\end{aligned}$$

Only the last quantity can drop below 1. ♠

Proof of Statement 7: We compute that the possible slopes for the distinguished edges are

$$\frac{1 + 2A - A^2}{A + A^2}, \quad \frac{-2A}{1 + A}, \quad \frac{1 - A}{1 + A}, \quad \frac{1 + 4A - A^2}{A^2 - 1}.$$

It is an easy exercise in algebra to show that these quantities avoid the set $\{-1, 0, 1, \infty\}$ for all $A \in (0, 1)$. ♠

13 Graph Master Picture Theorem

We will give a simplified and slightly modified account of the Graph Master Picture Theorem as it is stated in [S1, §6]. The changes we give here make the result easier to compare to the Plaid Master Picture Theorem. At the end of the chapter we discuss the modifications. In the three chapters following this one, we will prove a generalization of the Graph Master Picture Theorem which works for any polygon without parallel sides.

13.1 Statement of the Result

We fix $A = p/q$ as above. As in the case of the plaid PET, we work in the space $\widehat{X} = \mathbf{R}^3 \times [0, 1]$. This time, we use the coordinates (x, y, z, A) on \widehat{X} . The only difference is that we are calling the fourth coordinate A rather than P . The slices of the PET turn to be related by the equation $P = 2A/(1 + A)$, an equation we have encountered many times already.

Let Λ denote the abelian group of generated by the following affine transformations.

- $T_X(x, y, z, A) = (x + 1, y - 1, z - 1, A)$.
- $T_Y(x, y, z, A) = (x, y + 1 + A, z + 1 - A, A)$.
- $T_Z(x, y, z, A) = (x, y, z + 1 + A, A)$.

For each parameter A , the rectangular solid

$$R_A = [0, 1] \times [0, 1 + A] \times [0, 1 + A] \times \{A\} \quad (102)$$

serves as a fundamental domain for the action of Λ on $\mathbf{R}^3 \times \{A\}$. The quotient is some flat 3-torus which depends on A . The union

$$R = \bigcup_{A \in [0, 1]} R_A, \quad (103)$$

is a convex integer polytope.

We introduce a map $\Phi'_A : \mathbf{Z}^2 \rightarrow \widehat{X}$, as follows.

$$\Phi'_A(m, n) = (2t, 2t, 2t, A), \quad t = 2Am + 2n + \frac{1}{q}. \quad (104)$$

There are two Λ -invariant partitions of \widehat{X} into convex integral polytopes. Each polytope in each partition is labeled by a pair of integers $(i, j) \in \{-1, 0, 1\}$. The labeling has the property that the local structure of the arithmetic graph at the point $c \in \mathbf{Z}^2$ is determined by the labels of the polytopes in each partition which contain the image $\Phi'(c)$. In particular, $\Phi'(c)$ lies in the interior of a polytope in each partition.

Here is what we mean more precisely. Suppose that $\Phi'(c)$ lies in a polytope in the first partition with label (i_1, j_1) and a polytope in the second partition with labels (i_2, j_2) . Then the arithmetic graph has the edge connecting c to $c + (1_1, j_1)$ and the edge connecting c to $c + (i_2, j_2)$. We have $(i_1, j_1) = (0, 0)$ if and only if $(i_2, j_2) = (0, 0)$. In this case, c is an isolated point in the graph.

For each partition, a certain union of 14 polytopes forms a fundamental domain for the action of Λ . In both cases, the fundamental domain is an integral translate of the polytope R from Equation 103. The polytopes in the two partitions are related as follows. Define

$$I(x, y, z, A) = (1, A, 2 + A, A) - (x, y, z, 0). \quad (105)$$

Then

- $Q_j = I(P_j)$ for all $j = 1, \dots, 14$.
- The label of Q_j is the negative of the label of P_j .

The map I is an involution, so we have $I(Q_j) = P_j$ as well.

We call these two partitions the (+) graph partition and the (−) graph partition. We denote these partitions by \mathcal{G}_+ and \mathcal{G}_- . The *double partition* is $\mathcal{G}_+ \# \mathcal{G}_-$, the common refinement of the two partitions. Each polytope Z in the double partition is labeled by a quadruple (i_+, j_+, i_-, j_-) , where (i_\pm, j_\pm) is the label of the polytope in \mathcal{P}_\pm containing Z . The polytope of the double partition which contains $\Phi'_A(c)$ determines the two unoriented edges of the arithmetic graph incident to c . We do not need to know this partition explicitly.

In [S1, §6] we give a detailed geometric description of (translates of) the partitions \mathcal{G}_+ and \mathcal{G}_- . The geometric description is more intricate than what we did for the plaid PET. We will not repeat the description here. The reader can get a very clear picture of the partitions using our computer program. In §13.4 we list the 14 fundamental polytopes of \mathcal{G}_+ , together with their labels. As with the Plaid Master Picture Theorem, our listing, together with the formula for Φ'_A , gives a complete statement of the Graph Master Picture Theorem.

13.2 Pulling Back the Maps

As our notation suggests, we can improve on the map Φ'_A . Let T be the canonical affine transformation from §13. This map is given in Equation 98 and the inverse map is given in Equation 99. We can interpret the Master Picture Theorem as a statement about the structure of the affine image $T(\Gamma)$, where Γ is the arithmetic graph as defined in previous chapters. After this section, we will save words by using the term *arithmetic graph* to denote $T(\Gamma)$ rather than Γ . The vertices of $T(\Gamma)$ are contained in the graph grid $T(\mathbf{Z}^2)$.

To convert the Master Picture Theorem into a statement about the desired affine image, we simply pull back the classifying map. We define

$$\Phi_A = \Phi'_A \circ T^{-1}. \quad (106)$$

The domain for Φ is the graph grid G . A calculation shows that

$$\Phi(x, y) = (x, x, x, A). \quad (107)$$

This nice equation suggests that there is something canonical about the affine map that appears in the Quasi-Isomorphism Theorem. The graph grid G is more complicated than \mathbf{Z}^2 , but the classifying map Φ_A is much more natural.

13.3 Further Discussion

There are 3 small differences between our presentation of the Master Picture Theorem here and the one given in [S1]. For the reader who is interested in comparing what we say here to what we say in [S1], we discuss those differences.

Coordinate Swap: We have switched the first and third coordinates. Thus, the vertex (x, y, z, A) of a polytope here corresponds to the vertex (z, y, x, A) in [S1]. This change makes the Master Picture Theorem line up more gracefully with the Isomorphism Theorem for the plaid model. This switch also effects the definition of the lattice Λ .

Amalgamated Maps Here we have one classifying map and two partitions, whereas in [S1] we have two classifying maps, differing from each other by translations, and two slightly different partitions. The two partitions in [S1]

are translates of the ones here respectively by the vectors $(0, 1, 0, 0)$ and $(-1, 0, 0, 0)$. When these changes are made, the involution I simply becomes reflection in the midpoint of the fundamental domain R , and the union of the 14 fundamental polytopes, in either partition, is precisely R . In short, the partitions in [S1] are obtained from the ones here by translating so that everything lies in R . This picture is geometrically more appealing, but it serves our purposes here to have one map rather than two.

A Different Offset Even after we take into account the translations we just discussed, the map we have here is not quite the same as the map in [S1]. Were we to strictly translate the map from [S1] the small term $1/q$ in Equation 104 would be replaced by an infinitesimally small positive number $\iota = 0_+$. (If you like, you can interpret 0_+ as an infinitesimally small positive number in a nonstandard field containing \mathbf{R} .) Put another way, we set $\iota = 0$, and then in those rare cases when $p = \Phi'_A(m, n)$ lies in the boundary of one of the polytopes, we select the polytope which contains $p + (\epsilon, \epsilon, \epsilon, 0)$ for all sufficiently small ϵ . In fact, any choice of ι in $(0, 2/q)$ would give exactly the same result. Setting $\iota = 0$ and then stipulating the rules for handling boundary cases seemed at the time to simplify the picture. In hindsight, the choice $\iota = 1/q$, as we make here, is more canonical. Also, it is better suited to our present purposes. See Equation 107 in the next chapter.

Directed Graph: There is one more point we want to discuss. Since the arithmetic graph comes from following the dynamics of the first return map, it is possible to orient the paths in the arithmetic graph. It is also possible to determine when the first partition determines the edge pointing to c or the edge pointing away to c . This fine point is not needed for the proof of the Quasi-Isomorphism Theorem but it would be needed if we wanted to have a well-defined affine PET associated to the arithmetic graph, as we constructed for the plaid model.

More concretely, here is the recipe. We introduce the integer function

$$\rho_A(m, n) = \text{floor}(2s), \quad s = (2 + 2A)m + \frac{1}{q}. \quad (108)$$

Then \mathcal{G}_+ determines the outward pointing edge at (m, n) if and only if $\rho_A(m, n)$ is even. This result is implicit in the development given in [S1].

Our generalization of the Graph Master Picture Theorem, given in the next 3 chapters, automatically gives the PET structure.

13.4 The Fundamental Polytopes

Here are the 14 fundamental polytopes for the (+) partition. Note that the listing for the last one is spread over two lines.

$$\begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ +1 \\ +1 \end{bmatrix} \quad (0, +1)$$

$$\begin{bmatrix} 0 \\ -1 \\ 0 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ 0 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ 0 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ +1 \\ +1 \end{bmatrix} \quad (0, +1)$$

$$\begin{bmatrix} 0 \\ 0 \\ +1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ +1 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \\ +1 \end{bmatrix} \begin{bmatrix} 0 \\ +1 \\ 2 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ +1 \\ 2 \\ +1 \end{bmatrix} \quad (-1, 0)$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ +1 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ 2 \\ +1 \end{bmatrix} \quad (-1, 0)$$

$$\begin{bmatrix} 0 \\ -1 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ +1 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ 2 \\ +1 \end{bmatrix} \quad (+1, 0)$$

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ +1 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ +1 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ 2 \\ +1 \end{bmatrix} \quad (-1, -1)$$

$$\begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} +1 \\ +1 \\ 0 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ 0 \\ +1 \\ +1 \end{bmatrix} \begin{bmatrix} +1 \\ +1 \\ +1 \\ +1 \end{bmatrix} \quad (-1, 0)$$

14 Pinwheels and Quarter Turn Systems

14.1 Overview

In the next three chapters, we will prove a generalization of the Graph Master Picture Theorem which works for any convex polygon P without parallel sides. This material is not needed for the rest of the monograph. Throughout these three chapters, P denotes a convex polygon without any parallel sides. Let ψ denote the second iterate of the outer billiards map defined on $\mathbf{R}^2 - P$.

In this chapter we will define a map closely related to ψ , which we call the pinwheel map. For the purposes of studying unbounded orbits, the pinwheel map carries all the information contained in ψ . However, in general, the pinwheel map ignores “a bounded amount” of information contained in ψ . In the case of kites, the pinwheel map contains all the information. The general version of the Graph Master Picture Theorem compactifies the pinwheel map.

We will see that the pinwheel map has the structure of what we call a *quarter turn system*. A QTS is a certain kind of piecewise affine map of the infinite strip \mathbf{S} of width 1 centered on the x -axis. At the end of the chapter we will explicitly identify the QTSs which come from kites.

In §15, we will prove a general compactification result for a QTS, Theorem 15.1. We will see that this compactification can be encoded by a pair of lattices in \mathbf{R}^{n+1} and a pair of fundamental domains for those lattices. Here n is the number of sides of the polygon P . In the case of kites, the compactification lies in a 3-dimensional slice of \mathbf{R}^5 . We will discuss the case of kites at the end of §15.

In §16 we will prove the structural claims made at the end of §15.

The material in these three chapters is a compromise between giving no information about the Graph Master Picture Theorem and giving too much information. We already have a complete proof in [S1], so we thought that it would be nice here to give the theoretical framework behind the result. At the time I wrote [S1], I did not know this general framework. On the other hand, we do not explicitly translate the general result back to the special case presented in the last chapter. The translation would involve identifying the polytopes in the partition explicitly and then seeing that they match the listing given at the end of the last chapter.

14.2 The Pinwheel Map

Here we recall some work we did in [S2]. Let Σ be an infinite strip in the plane and let V be a vector that spans Σ in the sense that the tail of V lies on one component of $\partial\Sigma$ and the head of V lies on the other component. See Figure 14.1

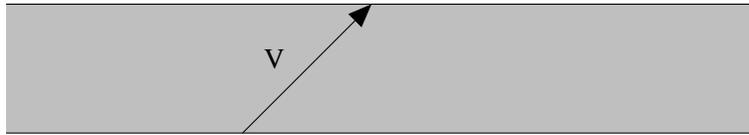


Figure 14.1: The vector V spans the strip.

The pair (Σ, V) defines a map $T : \mathbf{R}^2 \rightarrow \Sigma$, as follows.

$$T(p) = p + nV \tag{109}$$

Here $n \in \mathbf{Z}$ is the integer such that $p + nV \in \Sigma$. The map T is well-defined in the complement of a discrete infinite family of lines which are parallel to Σ . This family of lines contains the two lines of $\partial\Sigma$.

Let P and ψ be as above.

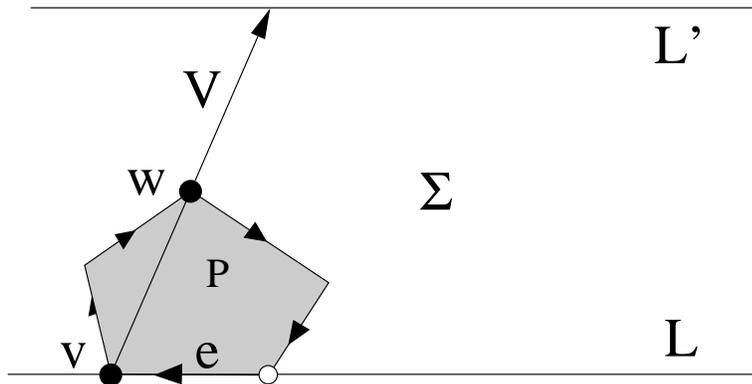


Figure 14.2: The strip associated to e .

We orient the edges of P clockwise. Given an edge e of P , we let L be the line extending e and we let L' be the line parallel to L so that the vertex w of P that lies farthest from L is equidistant from L and L' . We associate to e the pair (Σ, V) , where Σ is the strip bounded by L and L' , and $V = 2(w - v)$. See Figure 15.3

We order the strips according to their slopes, so that one turns counter clockwise when changing from Σ_i to Σ_{i+1} . This ordering typically does not coincide with the cyclic ordering on the edges. The corresponding composition

$$T = T_n \circ \dots \circ T_1 \tag{110}$$

is what we call *the pinwheel map*.

To describe the connection between T and outer billiards, we first work outside some large compact subset $K \subset \mathbf{R}^2$. Suppose we start with a point $p_1 \in \Sigma_1$. Then $\psi^k(p_1) = p_1 + kV_2$ for $k = 1, 2, 3, \dots$. This general rule continues until we reach an exponent k_1 such that $p_2 = \psi^{k_1}(p_1) \in \Sigma_2$. Then we have $\psi^k(p_2) = p_2 + kV_3$ for $k = 1, 2, 3, \dots$, until we reach an exponent k_2 such that $p_3 = \psi^{k_2}(p_2) \in \Sigma_3$. And so on. See Figure 1.3. We eventually reach a point $p_{n+1} \in \Sigma_1$, and the map $p_1 \rightarrow p_{n+1}$ is the first return map.

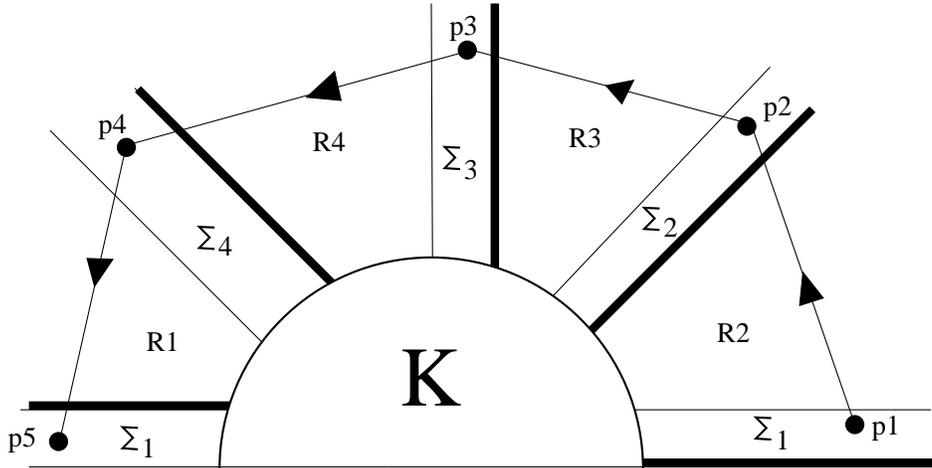


Figure 15.4: Far from the origin.

The connection between the first return map and ψ , for orbits *far away* from the polygon, appears in almost every paper on polygonal outer billiards. However, for points which start out near P , the connection is much less clear.

From the connection between outer billiards and the pinwheel map, we see one essential feature of our strips and vectors. (This connection can also be seen directly in terms of the polygon.) We have already mentioned that V_i spans Σ_i . Were we to consider the map ψ^{-1} in place of ψ , we would produce the sequence of strips $(\Sigma_{j-1}, -V_j)$. Therefore V_j spans Σ_{j-1} as well as Σ_j . We can put this in another way. $\Sigma_{j-1} \cap \Sigma_j$ is a parallelogram, and $2V_j$ is always one of the diagonals of this parallelogram.

14.3 Outer Billiards and the Pinwheel Map

Our main result in [S2] shows that the correspondence between the pinwheel map and outer billiards works well regardless of whether one starts near P or far away. Here we describe the main result in [S2]. We find it convenient to work with the square of the pinwheel map T^2 and the second return of ψ to Σ_1 , which we call Ψ^2 . The reason we do this is that both T^2 and Ψ^2 are (in a local sense) fairly close to the identity: They move points all the way around P rather than halfway around.

Theorem 14.1 *There is a disk $K \subset \mathbf{R}^2$ (depending on the polygon) with the following properties. Suppose that a, b are in the same component of $\Sigma_1 - K$. Then b is in the forward Ψ^2 orbit of a if and only if b is in the forward T^2 orbit of a .*

Suppose we start at a and iterate the two maps Ψ^2 and T^2 . Let's say we produce points c_1, c_2, \dots and c'_1, c'_2, \dots . As long as none of these points get close to the origin, we have $c_j = c'_j$ for all j and then we are simply saying that there is some index k so that $b = c_k = c'_k$. However, if the points move close to the origin. Then it might happen that $c_j \neq c'_j$. The orbits can diverge from each other. However, Theorem 14.1 is saying that these two sequences sync back up once they move far away from the origin. Here is a corollary of this result, also proved in [S2].

Theorem 14.2 *There is a canonical bijection between the set of unbounded orbits of ψ and the set of unbounded orbits of T . The bijection is such that the ψ -orbit O corresponds to the T -orbit which agrees with $O \cap \Sigma_1$ outside a compact set. In particular, outer billiards on P has unbounded orbits if and only if T has unbounded orbits.*

Theorem 14.2 tells us that we can replace the outer billiards system with the pinwheel map if we are only interested in the existence of unbounded orbits. However, for kites, we are interested in a more precise description of the outer billiards orbits. So, it might first appear that our switch to the pinwheel map causes us to lose some information about the orbits.

However, it turns out that for kites we have $T^2 = \Psi^2$ if we choose the domain carefully. (We will discuss this at the end of this chapter.) For this reason, our compactification for Ψ^2 in general gives rise to a compactification for T^2 when we are working with a kite. We will discuss the case of kites at the end of this chapter.

14.4 Quarter Turn Compositions

In this section and the next we analyze the structure of the pinwheel map. Let \mathbf{S} denote the strip of width 1 whose centerline is the x -axis.

Let $\square \subset \mathbf{S}$ be a rectangle with sides parallel to the coordinate axes. The top and bottom of \square are supposed to lie in the top and bottom boundary of \mathbf{S} . We define a *quarter turn* of \square to be the order 4 affine automorphism of \square which maps the right edge of \square to the bottom edge of \square . This map essentially twirls \square one quarter of a turn clockwise. For any $a > 0$ we distinguish 2 tilings of the strip \mathbf{S} by $a \times 1$ rectangles. In *Tiling 0*, the origin is the center of a rectangle. In *Tiling 1*, the origin is the center of a vertical edge of a rectangle. For $q = 0, 1$ let $R_{q,a}$ denote the map which gives a quarter turn to each rectangle in Tiling q . The map $R_{q,a}$ is a piecewise affine automorphism of \mathbf{S} , defined everywhere except the vertical edges of the rectangles. We call $R_{q,a}$ a *quarter turn*.

We define the *shear*

$$S_s = \begin{bmatrix} 1 & -s \\ 0 & 1 \end{bmatrix} \quad (111)$$

Here $s > 0$. The map S_s is a shear of \mathbf{S} which fixes the centerline pointwise, moves points with positive y -coordinate backwards and points with negative y -coordinate forwards.

We define a *quarter turn composition* (QTS) to be a finite alternating composition \mathcal{T} of quarter turns and shears. That is,

$$\mathcal{T} = S_{s_n} \circ R_{q_n, r_n} \circ \cdots \circ S_{s_1} \circ R_{q_1, r_1}. \quad (112)$$

- $q_1, \dots, q_n \in \{0, 1\}$ specify the tiling offsets.
- r_1, \dots, r_n are the parameters for the widths of the rectangles.
- $s_1, \dots, s_n > 0$ are the parameters for the shears.

We call n the *length* of the QTS.

It is convenient to define

$$\alpha_i = r_n / r_i \quad (113)$$

The choice of n as a special index is arbitrary; any other choice leads to the same definitions. We call \mathcal{T} *quasi-rational* if $\alpha_i \in \mathbf{Q}$ for all i .

We call \mathcal{T} *finitary* if \mathcal{T} is a piecewise translation, and the set

$$\{\mathcal{T}(p) - p \mid p \in \mathbf{S}\} \quad (114)$$

of possible translations is finite.

14.5 The Pinwheel Map as a QTS

Now we will recognize the pinwheel map as a QTS. We find it useful to consider the map T^2 rather than the map T because the map T^2 is fairly close to the identity map; it corresponds to moving the point all the way around P rather than halfway around.

Let $(\Sigma_1, V_1), \dots, (\Sigma_{2n}, V_{2n})$ denote the strip data above, repeated twice. Let T_1, \dots, T_{2n} be the corresponding strip maps. For the purpose of getting the signs right when we define certain maps, we normalize the picture by a suitable affine transformation. We assume that $\Sigma_1 = \mathbf{S}$ and that $\Sigma_2, \dots, \Sigma_n$ all have positive slope. However, after we define our maps, we will not insist on this normalization. For $k = 1, \dots, n$, we define map

$$A_{k,\pm} : \Sigma_k \rightarrow \mathbf{S}, \quad (115)$$

by the following rules.

- $A_{k,\pm}$ is area preserving, orientation preserving, and maps points with large positive x -coordinate to points with large positive x -coordinate.
- $A_{k,\pm}$ maps the parallelogram $\Sigma_k \cap \Sigma_{k\pm 1}$ to a rectangle, and the head point of V_k to the origin.

Let $\rho(x, y) = (-x, -y)$ be reflection about the origin. We define

$$A_{n+k,\pm} = \rho \circ A_{k,\pm}, \quad (116)$$

$$R_k = A_{k+1,-} \circ T_k \circ A_{k,+}^{-1}; \quad S_k = A_{k+1,+} \circ (A_{k+1,-})^{-1} \quad (117)$$

Now that we have defined these maps, we drop the assumption about the slopes of the strips. In general, the maps are defined in such a way that the whole construction is natural under affine conjugation.

Lemma 14.3 *S_k is an affine shear, as in Equation 111.*

Proof: The maps $A_{k+1,\pm}$ are both area preserving, orientation preserving, sense preserving affine bijections from Σ_{k+1} to \mathbf{S} , and they both map the same point to the origin. From this description, it is clear that S_k has the equation given in Equation 111. The only thing that remains to prove that $b > 0$. That is, S_k shears points in \mathcal{S} with positive y -coordinate to the left.

To understand what is going on, we take $k = 1$. Now we normalize so that Σ_1 is horizontal and Σ_2 is vertical, and the positive senses of these strips go along the positive coordinate axes. See Figure 15.5. This situation forces Σ_3 to have negative slope. $\Sigma_1 \cap \Sigma_2$ is the thickly drawn square and $\Sigma_2 \cap \Sigma_3$ is the shaded parallelogram. The significant feature here is that the left side of the shaded parallelogram lies above the right side. This fact translates into the statement that $b > 0$ in Equation 111. ♠

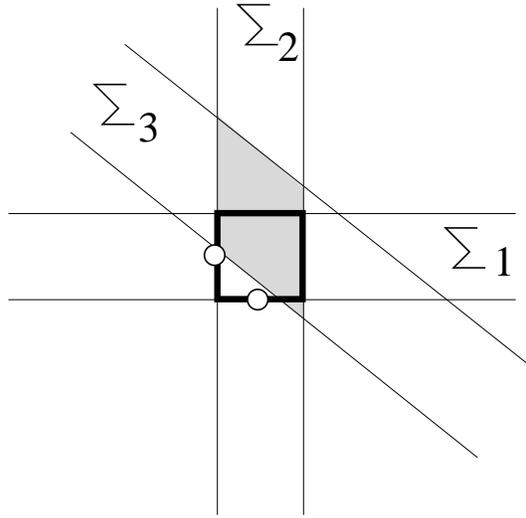


Figure 15.5: Placement of the strips

Lemma 14.4 R_k is a quarter turn map.

Proof: Our proof will also identify the parameters of R_k . The parameter r_k is just the area of $\Sigma_k \cap \Sigma_{k-1}$. We will consider the case $k = 1$. We first discuss how our construction interacts with affine transformations. Let Δ be an affine transformation, which expands areas by δ . Let R'_1 be the map associated to $\Delta(P)$. We have

$$R'_1 = D_\delta \circ R_1 \circ D_\delta^{-1}, \quad D_\delta(x, y) = (\delta x, y).$$

Thanks to this equation, it suffices to prove our result for any affine image of P . We normalize by an affine transformation so that

$$\Sigma_1 = \mathbf{R} \times [-1/2, 1/2], \quad \Sigma_2 = [-1/2, 1/2] \times \mathbf{R}, \quad V_2 = (-1, 1). \quad (118)$$

In this case, $A_{1,+}$ is the identity and $A_{2,-}$ is the clockwise order 4 rotation about the origin. Figure 15.6 shows the action of T_2 on Σ_1 .

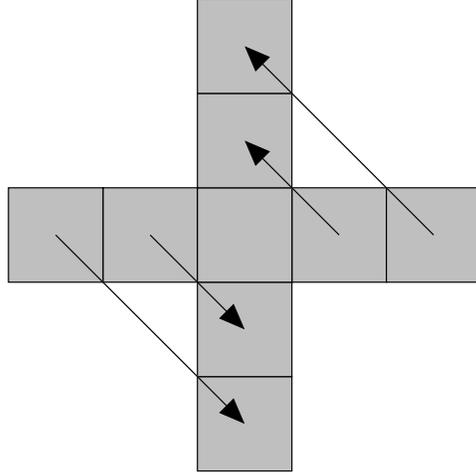


Figure 15.6: Action of T_2 .

From Figure 15.6, and from the description of $A_{1,+}$ and $A_{2,-}$, we see that there is a tiling \mathcal{S} of \mathbf{S} by unit squares and R_1 gives a clockwise quarter turn to each unit square. To finish the proof, we just have to see that \mathcal{S} is one of the two special tilings discussed in §14.4.

Let e_j be the edge of P that lies in $\partial\Sigma_j$. Let $|V_j|$ is the segment underlying the vector V_j . One basic principle we use in our analysis is that $e_1, |V_1|, |V_2|$ make the edges of a triangle. Call this the *triangle property*.

Let c_j be the head of V_j . We have

$$R_{1,+}(c_1) = (0,0) = R_{2,-}(c_2). \quad (119)$$

There are two cases to consider. Suppose that c_1 is not incident to e_2 . By the triangle property, c_1 is incident to V_2 . Hence either $c_1 = c_2$ or c_1 is the tail vertex of V_2 . But the tail vertex of V_2 is incident to e_2 . This proves that $c_1 = c_2$. This situation implies that the common point $c = c_1 = c_2$ is the center of $\Sigma_1 \cap \Sigma_2$. From this information, and Equation 119, we conclude that \mathcal{S} is Tiling 1. Suppose that c_1 is incident to e_2 . Then c_1 lies on the centerline of Σ_1 and on the boundary of Σ_2 . Hence c_1 is the midpoint of an edge of $\Sigma_1 \cap \Sigma_2$. Hence, the origin is the midpoint of an edge of a tile in \mathcal{S} . Hence \mathcal{S} is Tiling 2. ♠

We define

$$\mathcal{T}_P = S_n \circ R_n \circ \cdots \circ S_1 \circ R_1. \quad (120)$$

By construction, \mathcal{T}_P is a QTS.

Let T be the pinwheel map. By construction

$$T = \rho \circ \mathcal{T}_P; \quad T^2 = S_{2n} \circ R_{2n} \circ \cdots \circ S_1 \circ R_1. \quad (121)$$

Lemma 14.5 \mathcal{T}_P^2 is finitary.

Proof: We have $\mathcal{T}_P^2 = T^2$. The map T^2 is evidently a piecewise translation. We just need to prove that the set $\{T^2(p) - p \mid p \in S\}$ is finite. There is a sequence of numbers m_1, \dots, m_{2n} such that

$$T^2(p) - p = \sum_{i=1}^{2n} m_i V_i = \sum_{i=1}^n (m_i - m_{i+n}) V_i$$

Here V_1, \dots, V_n are the vectors that arise in the strip maps, and we are setting $V_{i+n} = -V_i$. The number m_j refers to the analysis of Figure 15.4. Here m_j is the number of iterates of ψ needed to carry the iterate lying in Σ_{j-1} to the iterate lying in Σ_j .

Far from the origin, the portion of the ψ -orbit of p , going from p to $T^2(p)$, lies within a uniformly bounded distance of a centrally symmetric $2n$ -gon. The point is that all the strips come within a uniform distance of the origin. From this property, we see that there is a uniform bound to $|m_i - m_{i+n}|$ for all i . Hence, there are only finitely many choices for $T^2(p) - p$. ♠

14.6 The Case of Kites

Now we explain the situation for the kite K_A . Define

$$B = \frac{1+A}{1-A}, \quad C = \frac{B^2}{C^2-1}. \quad (122)$$

The QTS corresponding to K_A has parameters

$$q = (1, 1, 0, 1) \quad r = (1, B, C, B); \quad s = (B, BC, BC, B). \quad (123)$$

The special orbits on K_A correspond to the points on the thick diagonal lines drawn in Figure 15.7. These diagonal lines are invariant under the QTS.

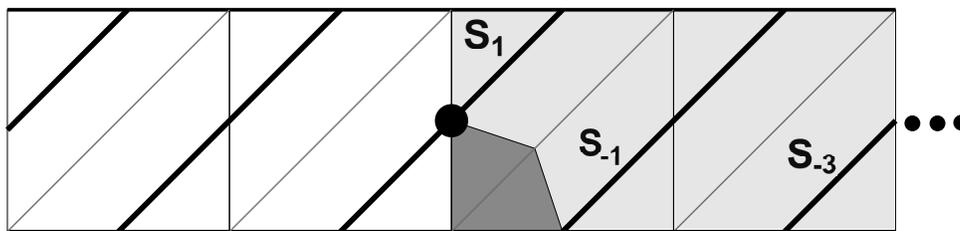


Figure 15.7: The set I_1 contained in \mathbf{S} .

In the case of K_A , the connection between outer billiards and the pinwheel map is especially tight. Let Σ_1^+ denote the portion of Σ_1 which lies to the right of the edge of K_A having slope 1. The corresponding subset of \mathbf{S} is lightly shaded in Figure 15.7. On Σ_1^+ , we have $\Psi^2 = T^2$. The pinwheel map and the first return map of ψ to Σ_1^+ coincide.

There are two transformations we need to make to get the picture in Figure 15.7 back to the picture we considered when we defined the arithmetic graph. First, in order to get Σ_1 from \mathbf{S} we need to apply the similarity h which expands by $2\sqrt{2}$, rotates clockwise by $\pi/2$, and maps $(0, 0)$ to $(0, 1)$. The map hT^2h^{-1} coincides with the first return map of ψ to Σ_1^+ . The map h carries the thick diagonal lines in Figure 15.7 to the intersection of Σ_1^+ with the horizontal lines having odd integer y -intercept. Call this set \implies' .

Now we describe the second transformation. For the purposes of defining the arithmetic graph, we are not quite interested in the first return map of ψ to \implies' . Rather, we are interested in the first return map of ψ to \implies , the two horizontal rays emanating from $(0, \pm 1)$ and pointing along the positive x -axis. However, the powers of ψ give a bijection between \implies' and \implies which is continuous at all points where we have well-defined orbits. The bijection just amounts to adding various multiples of the vector $(0, 4)$ to points of \implies' .

You can see this second transformation in Figure 15.7. The segment $S_{-3} + (-1, 1)$ continues the segment S_1 . In general the diagonal ray emanating from the top vertex of the shaded kite is the union

$$S_1 \cup (S_{-3} + (-1, 1)) \cup (S_{-5} + (-2, -2)) \cup \dots$$

When you apply h to this picture you see how ψ sets up a (near) bijection between half of \implies' and the top ray of \implies . The picture for the bottom ray is similar.

15 The General Compactification

15.1 Affine Pets Redefined

In this chapter we will prove a general compactification theorem for quarter turn compositions. Define the unit torus

$$\widehat{\mathbf{S}} = \mathbf{R}^{n+1} / \mathbf{Z}^{n+1}. \quad (124)$$

It is convenient to let $\mathbf{T}^d = \mathbf{R}^d / \mathbf{Z}^d$ denote the unit torus made from the first d -coordinates of \mathbf{R}^{n+1} . Thus $\widehat{\mathbf{S}} = \mathbf{T}^{n+1}$. Here is the main result of this chapter.

Theorem 15.1 *Suppose that \mathcal{T} is a length- n quarter turn composition. Then there is a locally affine map $\Psi : \mathbf{S} \rightarrow \widehat{\mathbf{S}}$ and an affine PET, $\widehat{\mathcal{T}} : \widehat{\mathbf{S}} \rightarrow \widehat{\mathbf{S}}$, such that $\Psi \circ \mathcal{T} = \widehat{\mathcal{T}} \circ \Psi$.*

- *The map Ψ is injective if and only if \mathcal{T} is not quasi-rational.*
- *The closure of $\Psi(\mathbf{S})$ is a sub-torus of dimension $1 + d$, where d is the \mathbf{Q} -rank of $\mathbf{Q}(\alpha_1, \dots, \alpha_{n-1})$.*
- *If \mathcal{T}^k is finitary, then the restriction of $(\widehat{\mathcal{T}})^k$ to the closure of $\Psi(\mathbf{S})$ is an ordinary PET.*

We will see that when \mathcal{T} comes from the pinwheel map, as in the previous chapter, the compactification $\widetilde{\mathcal{T}}^2$ is an ordinary PET. When we specialize to the case of kites, and massage the result a little, we will get the Graph Master Picture Theorem.

For the purposes of proving Theorem 15.1, we do not insist that the ambient space is a convex polytope. Rather, we will take the ambient space to be a flat torus. In the next chapter we will recognize the compactification as an affine PET in the sense of the introduction.

We say that a dense open set $U \subset \widehat{\mathbf{S}}$ is an *invariant domain* for an affine PET f if f is entirely defined on U and $f(U) = U$. When it comes time to recognize the affine PETs from Theorem 15.1 as affine PETs in the sense defined in the introduction, we will find an invariant domain that is isometric to a parallelotope.

15.2 The Map

Let \mathcal{T} , \mathbf{S} and $\widehat{\mathbf{S}}$ be as in Theorem 15.1. We define $\Psi : \mathbf{S} \rightarrow \widehat{\mathbf{S}}$ by the formula

$$\Psi(x, y) = (\psi(x), [y]); \quad \psi(x) = \left[\frac{x}{r_1}, \dots, \frac{x}{r_n} \right]. \quad (125)$$

Here $[p]$ denotes the image of the point p in the relevant space $(\mathbf{R}/\mathbf{Z})^k$. So, for instance, $[y] \in \mathbf{R}/\mathbf{Z}$.

Note that Ψ is injective if and only if \mathcal{T} is not quasi-rational. For kites, \mathcal{T} is quasi-rational if and only if the kite is rational. Let \mathbf{S}^* denote the closure of the image of $\Psi(\xi)$.

Lemma 15.2 $\dim(\mathbf{S}^*) = d + 1$, where d is the dimension of $\mathbf{Q}(\alpha_1, \dots, \alpha_{n-1})$.

Proof: It suffices to prove that the closure X of $\psi(\mathbf{R})$ in \mathbf{T}^n has dimension d . Permuting the coordinates, it suffices to consider the case when $\alpha_{n-d+1}, \dots, \alpha_{n-1}, 1$ are independent over \mathbf{Q} and α_j is a rational combination these last d variables for all $j \leq n - d$. Let $\pi : \mathbf{T}^n \rightarrow \mathbf{T}^d$ be projection onto the last d coordinates. By the previous result, $\pi(X) = \mathbf{T}^d$. To prove that $\dim(X) = d$ it suffices to prove that $X \cap \pi^{-1}(0, \dots, 0)$ consists of finitely many points. Let p be a point in this intersection. We will show that the first coordinate of p can only take on finitely many values. The same argument works for the remaining coordinates.

We have some integer relation

$$c_1 \alpha_1 = c_{n-d+1} \alpha_{n-d+1} + \dots + c_{n-1} \alpha_{n-1} + c_n. \quad (126)$$

Multiplying through by r_n we have

$$\frac{c_1}{r_1} = \frac{c_{n-d+1}}{r_{n-d+1}} + \dots + \frac{c_n}{r_n}. \quad (127)$$

Suppose $x \in \mathbf{R}$ is such that $\pi \circ \psi(x)$ is close to $(0, \dots, 0)$. Then x/r_j is close to an integer for $j = n - d + 1, \dots, n$. But then $c_j x/r_j$ is also close to an integer for $j = n - d + 1, \dots, n$. But then $c_1 x/r_1$ is close to an integer. This argument shows that the first coordinate of any point of $F \cap \pi^{-1}(0, \dots, 0)$ has the form $[k/c_1]$ for some $k \in \{1, \dots, c_1\}$. In particular, this is a finite set of possibilities. ♠

15.3 Extending The Component Maps

A QTS is the composition of two kinds of maps. In this section we treat each of these maps in isolation.

Lemma 15.3 *Let S be a shear of \mathbf{S} in Equation 111. There is an affine PET $\Psi \circ S = \widehat{S} \circ \Psi$, where*

$$\widehat{S}([x_1, \dots, x_n, y]) = \left[x_1 - \frac{s}{r_1}y, \dots, x_n - \frac{s}{r_n}y, y \right]. \quad (128)$$

\widehat{S} is an affine PET, and $\mathbf{T}^n \times (-1/2, 1/2) \subset \widehat{\mathbf{S}}$ is an invariant domain for \widehat{S} . The linear part of \widehat{S} is given by the matrix

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & -s/r_1 \\ 0 & 1 & 0 & \cdots & -s/r_2 \\ 0 & 0 & 1 & \cdots & -s/r_3 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (129)$$

Proof: We have $S(x, y) = (x - sy, y)$. A direct calculation shows that $\Psi \circ S = \widehat{S} \circ \Psi$ for the map \widehat{S} given above. Once we have the map \widehat{S} , the given domain is clearly an invariant domain. ♠

Lemma 15.4 *Let $R_n = R_{q_n, r_n}$. There exists an affine PET $\widehat{R}_n : \widehat{\mathbf{S}} \rightarrow \widehat{\mathbf{S}}$ such that $\Psi \circ R_n = \widehat{R}_n \circ \Psi$. The linear part of \widehat{R}_n is given by the matrix*

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & -r_n/r_1 & r_n/r_1 \\ 0 & 1 & 0 & \cdots & -r_n/r_2 & r_n/r_2 \\ 0 & 0 & 1 & \cdots & -r_n/r_3 & r_n/r_3 \\ \cdots & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix} \quad (130)$$

Define

$$X_{0,n} = \mathbf{T}^{n-1} \times (-1/2, 1/2) \times (-1/2, 1/2); \quad X_{1,n} = \mathbf{T}^{n-1} \times (0, 1) \times (-1/2, 1/2) \quad (131)$$

When $q_n = k$, the set $X_{k,n}$ is an invariant domain for \widehat{R}_n .

Proof: Suppose $q_n = 0$. Let

$$\Psi^*(x, y) = \left[\frac{x^*}{r_1}, \dots, \frac{x^*}{r_{n-1}}, \frac{x}{r_n}, y \right], \quad x^* = r_n \operatorname{int}\left(\frac{x}{r_n}\right).$$

Here $\operatorname{int}(x)$ is the integer nearest x . If $(x_2, y_2) = R_n(x_1, y_1)$, then $x_2^* = x_1^*$. Hence $\Psi^* \circ R_n = F \circ \Psi^*$, where F does nothing to the first $n - 1$ coordinates and, with respect to the last two coordinates, acts as an order 4 clockwise rotation fixing $[0, 0]$. That is,

$$F([x_1, \dots, x_{n-1}, x_n, y]) = [x_1, \dots, x_{n-1}, y + (q_n/2), -x_n + (q_n/2)]. \quad (132)$$

On $X_{0,n}$, we have $\Psi = Y \circ \Psi^*$, where

$$Y([x_1, \dots, x_{n-1}, x_n, y]) = \left[x_1 + \frac{r_n}{r_1} x_n, \dots, x_{n-1} + \frac{r_n}{r_{n-1}} x_n, x_n, y \right]. \quad (133)$$

The set $X_{0,n}$ evidently is an invariant domain for both Y and F . The map

$$\widehat{R}_n = Y \circ F \circ Y^{-1}, \quad (134)$$

has all the desired properties. A short exercise in matrix multiplication shows that the linear part of \widehat{R}_n has the form given in Equation 130.

When $q_n = 1$ we use the floor function in place of the nearest integer function when defining Ψ^* . ♠

The same result as above holds for R_{q_j, n_j} . The only difference is that the roles played by the indices j and n are swapped. For instance, the linear part of \widehat{R}_1 is given by the matrix

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ -r_1/r_2 & 1 & 0 & \cdots & 0 & r_1/r_2 \\ -r_1/r_3 & 0 & 1 & \cdots & 0 & r_1/r_3 \\ \cdots & & & & & \\ -r_1/r_n & 0 & 0 & \cdots & 1 & r_1/r_n \\ -1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}, \quad (135)$$

and the invariant domain is obtained from one of the domains in Equation 131 by permuting the 1st and n th coordinates.

15.4 The Composition

We now define

$$\widehat{\mathcal{T}} = \widehat{S}_{s_n} \circ \widehat{R}_{q_n, r_n} \circ \cdots \circ \widehat{S}_{s_1} \circ \widehat{R}_{q_1, r_1}. \quad (136)$$

The composition of affine pets is an affine PET. so $\widehat{\mathcal{T}}$ is an affine PET. By construction, $\Psi \circ \mathcal{T} = \widehat{\mathcal{T}} \circ \Psi$.

Let \mathbf{S}^* denote the closure of $\Psi(\mathbf{S})$ in $\widehat{\mathbf{S}}$. Now we suppose that \mathcal{T}^k is finitary for some exponent k . We will prove that the restriction of $\widehat{\mathcal{T}}^k$ to \mathbf{S}^* is an ordinary PET. For ease of notation, we assume that $k = 1$. The proof works the same way regardless of exponent.

We just need to show that $\widehat{\mathcal{T}}$ is a local translation. Suppose that $\widehat{p} \in \mathbf{S}^*$ and $\{\widehat{p}_n\}$ is a sequence of points in \mathbf{S}^* converging to \widehat{p} . We want to show that

$$\widehat{\mathcal{T}}(\widehat{p}) - p = \widehat{\mathcal{T}}(\widehat{p}_n) - p_n, \quad (137)$$

for all n sufficiently large. Since $\Psi(\mathbf{S})$ is dense in \mathbf{S}^* and the linear part of $\widehat{\mathcal{T}}$ is independent of point, it suffices to consider the case when $\widehat{p} = \Psi(p)$ and $\widehat{p}_n = \Psi(p_n)$ for some $p \in \mathbf{S}$ and some sequence $\{p_n\}$ in \mathbf{S} . Note that $\{p_n\}$ need not be a convergent sequence in \mathbf{S} .

Lemma 15.5 *Setting $V_s = \mathcal{T}(s) - s$ for any $s \in S$, we have*

$$\widehat{\mathcal{T}}(p) - p = \Psi(V_p), \quad \widehat{\mathcal{T}}(p_n) - p_n = \Psi(V_{p_n}). \quad (138)$$

Proof: We have

$$\widehat{\mathcal{T}}(\widehat{p}) - \widehat{p} = \widehat{\mathcal{T}} \circ \Psi(p) - \Psi(p) = \Psi \circ \mathcal{T}(p) - \Psi(p) = \Psi(V_p) \quad (139)$$

The last equality comes from the fact that $\Psi(V+W) = \Psi(V) + \Psi(W)$ whenever V, W , and $V+W$ all belong to S . Here we are taking $V = V_p$ and $W = p$. The same argument works for p_n . ♠

We now observe the following properties.

1. By continuity, $\Psi(V_{p_n}) \rightarrow \Psi(V_p)$ as $n \rightarrow \infty$.
2. Since \mathcal{T} is finitary, there is a uniform upper bound to $|V_{p_n}|$.
3. Ψ is injective.

It follows from these properties that $V_{p_n} \rightarrow V_p$. But \mathcal{T} is finitary. Hence $V_{p_n} = V_p$ for n large. But then $\Psi(V_p) = \Psi(V_{p_n})$ for n large. This fact combines with Equation 138 to establish Equation 137 for n large.

15.5 Double Lattice PETs

Now we explain the structure of the compactification from Theorem 15.1.

We say that a *double lattice PET* is a pair $(X_1, \Lambda_1, X_2, \Lambda_2)$, where

- X_1 and X_2 are parallelotopes in \mathbf{R}^{n+1} .
- Λ_1 and Λ_2 are lattices in \mathbf{R}^{n+1} .
- X_i is a fundamental domain for Λ_j for all $i, j \in \{1, 2\}$.

At first it might seem very difficult to find such objects, but Theorem 15.1 produces them in abundance.

Let $\phi : X_1 \cup X_2 \rightarrow X_1 \cup X_2$ be the map with the following definition.

- If $p \in X_1 \cap X_2$, then $\phi(p) = p$.
- If $p \in X_1 - X_2$ then $\phi(p) = (p + \Lambda_2) \cap X_2$.
- If $p \in X_2 - X_1$ then $\phi(p) = (p + \Lambda_1) \cap X_1$.

The partition for ϕ is

$$(\Lambda_1 X_1 \# \Lambda_2 X_2)|_{X_1 \cup X_2}. \quad (140)$$

Here $\Lambda_j X_j$ is the partition of \mathbf{R}^{n+1} by Λ_j translates of X_j . The symbol ($\#$) denotes the common refinement of the partitions. The inverse map ϕ^{-1} has the same construction, except with the roles of Λ_1 and Λ_2 reversed. So, the partition for ϕ^{-1} is

$$(\Lambda_2 X_1 \# \Lambda_1 X_2)|_{X_1 \cup X_2}. \quad (141)$$

The map ϕ^2 preserves both X_1 and X_2 separately, so if we want to get an ordinary PET with a convex domain, we can restrict ϕ^2 to X_1 . Of course, there are 3 other PETs we can get with the same data. All in all, we can restrict $\phi^{\pm 2}$ to X_j for $j \in \{1, 2\}$. All these PETs have essentially the same dynamics.

We want to recognize the compactification from Theorem 15.1 as a double lattice PET, at least in the case when the QTC comes from a pinwheel map. This is not quite the case. We will get a slightly different kind of PET, but then we will be able to make a straightforward change of coordinates to arrive at the double lattice PET. Here we describe the variant of the double lattice PET which we actually encounter.

15.6 The Structure Theorem

One can specify an affine PET by giving a triple (X_1, X_2, I) , where X_1 and X_2 are polyhedral fundamental domains for \mathbf{Z}^{n+1} and I is a linear isomorphism from X_1 to X_2 . The affine PET is given by

$$[X_1, X_2, I] := \Pi_2 \circ I \circ \Pi_1^{-1} \quad (142)$$

Here Π_j is the canonical map from X_j to $\widehat{\mathbf{S}}$. The map $\Pi_j^{-1} : \widehat{\mathbf{S}} \rightarrow X_j$ is defined as follows: Lift to \mathbf{R}^{n+1} then translate by the appropriate integer vector. When I is an involution, $[X_1, X_2, I]^2$ is an ordinary PET. In §16 we prove the following result.

Theorem 15.6 *The affine PET from Theorem 15.1 is conjugate to a map $[X_1, X_2, I]$, where X_1 and X_2 are parallelotope fundamental domains for \mathbf{Z}^{n+1} , centered at the origin. The map I fixes pointwise a codimension 2 subspace of \mathbf{R}^{n+1} and preserves each 2-plane parallel to $\Psi(\mathbf{S})$.*

When our QTC comes from a pinwheel map, I is an involution and indeed has eigenvalues $-1, -1, 1, \dots, 1$. In this case we can recover the double lattice PET structure. Consider the map

$$\Psi' = \Pi_1^{-1} \circ \Psi : \mathbf{S} \rightarrow X_1. \quad (143)$$

The map Ψ' is a piecewise affine map. After a bit of algebra, we get

$$\Psi' \circ \mathcal{T} = (F \circ I) \circ \Psi', \quad F = \Pi_1^{-1} \circ \Pi_2. \quad (144)$$

The map F has a very simple description. Given a generic $p \in X_2$ we have $F(p) = p + v$, where $v \in \mathbf{Z}^{n+1}$ is the unique vector such that $p + v \in X_1$.

Let $\Lambda_1 = \mathbf{Z}^{n+1}$ and $\Lambda_2 = I(\mathbf{Z}^{n+1})$. Note that X_i is a fundamental domain for Λ_j for all pairs (i, j) and the involution I has the action $\Lambda_1 \leftrightarrow \Lambda_2$ and $X_1 \leftrightarrow X_2$.

ϕ commutes with I and

$$\Psi' \circ \mathcal{T} = (I \circ \phi)|_{X_1} \circ \Psi', \quad \Psi' \circ \mathcal{T}^2 = \phi^2|_{X_1} \circ \Psi'. \quad (145)$$

The map ϕ^2 is exactly the map that we associated to the double lattice PET $(X_1, X_2, \Lambda_1, \Lambda_2)$. If we are willing to use the map Ψ' , we see that the second iterate of a QTC compactified by a double lattice PET when the QTC comes from a pinwheel map.

15.7 The General Arithmetic Graph

There is an arithmetic graph associated to the triple $[X_1, X_2, I]$. The vertices of this graph are the points of \mathbf{Z}^{n+1} . If $[X_1, X_2, I]^2$ is defined on $p \in \widehat{\mathbf{S}}$ then there is a unique vector $v_p \in \mathbf{Z}^{n+1}$ so that $\Pi_1^{-1}(p) + I(v_p) \in X_2$. This derives from the fact that both X_1 and X_2 are fundamental domains for \mathbf{Z}^{n+1} . A calculation shows that

$$[X_1, X_2, I]^2(p) = p + I(v_p). \quad (146)$$

The addition takes place in $\widehat{\mathbf{S}}$. It turns out that the last coordinate of v_p is 0.

Given an orbit $O = \{p_k\}$ we define Γ_O to be the lattice path in \mathbf{Z}^{n+1} such that

$$\Gamma_O(m+1) - \Gamma_O(m) = v_{p_m}. \quad (147)$$

We call Γ_O the *arithmetic graph* of the orbit. Γ_O is only defined up to integer translation.

There is another way to think about the arithmetic graph, in which we consider many orbits of the same time. Our construction depends on the choice of an *offset vector* $V_0 \in \mathbf{R}^{n+1}$. Define the set

$$S(V_0) = \{V_0 + I(\mathbf{Z}^{n+1})\} \pmod{\mathbf{Z}^{n+1}}. \quad (148)$$

The countable set $S(V_0) \subset \widehat{\mathbf{S}}$ is invariant under the action of $[X_1, X_2, I]^2$. That is, $S(V_0)$ is partitioned into orbits of $[X_1, X_2, I]^2$.

Define $\mu : \mathbf{Z}^{n+1} \rightarrow S(V_0)$ by the equation

$$\mu(V) = I(V) \pmod{\mathbf{Z}^{n+1}}. \quad (149)$$

Given $V, V' \in \mathbf{Z}^{n+1}$, we join V and V' by a directed edge if and only if

$$\mu(V') = [X_1, X_2, I]^2(\mu(V)). \quad (150)$$

This construction produces a directed graph $\Gamma(V_0)$ whose vertex set is \mathbf{Z}^{n+1} . Each component of $\Gamma(V_0)$ is the arithmetic graph of some orbit of $S(V_0)$, and the arithmetic graph of every such orbit arises as a component.

Note that changing the offset vector V_0 will likely produce a different graph. However, in practice, one can get a general sense of what the graph is like just by picking some random V_0 and drawing pictures.

15.8 The Case of Kites

Now we discuss how Theorems 15.1 and 15.6 work for the kite K_A . The map $\Psi : \mathbf{S} \rightarrow \widehat{\mathbf{S}}$ is given by

$$\Psi(x, y) = \left(x, \frac{1-a}{1+a}x, \frac{4a}{(1+a)^2}x, \frac{1-a}{1+a}x, y \right) \pmod{\mathbf{Z}^5}. \quad (151)$$

The compactification of \mathcal{T} is typically 4 dimensional. In view of the fact that $r_2 = r_4$, this compactification is contained in

$$(\mathbf{R}^5/\mathbf{Z}^5) \cap \{x_2 = x_4\}. \quad (152)$$

Letting Ξ' be the diagonal line segments discussed in §14.6, we note that $\Psi(\Xi')$ is contained in an invariant subspace given by

$$\{x_2 = x_4\} \cap \{x_1 = y\}. \quad (153)$$

The corresponding PET is typically 3-dimensional. One can derive the partition from the Graph Master Picture Theorem from this PET.

Now we describe the associated triple (X_1, X_2, I) from Theorem 15.6. The involution I is given by

$$I = \begin{bmatrix} 0 & -1 & \frac{-1-a}{2} & 0 & 0 \\ \frac{a-1}{a+1} & \frac{2a}{1+a} & \frac{a-1}{2} & 0 & 0 \\ \frac{-4a}{(1+a)^2} & \frac{-4a}{(1+a)^2} & \frac{1-a}{1+a} & 0 & 0 \\ \frac{a-1}{a+1} & \frac{a-1}{a+1} & \frac{a-1}{2} & 1 & 0 \\ 1 & 0 & \frac{-1-a}{2} & -1 & -1 \end{bmatrix} \quad (154)$$

X_j consists of those vectors v such that $M_j(v) \in [-1/2, 1/2]^5$. Here M_1 and M_2 respectively are

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{2a}{1+a} & 1 & 0 & 0 & \frac{1-a}{1+a} \\ \frac{-2a}{1+a} & \frac{1-a}{1+a} & 1 & 0 & 1 \\ -1 & 0 & \frac{1+a}{2} & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & -1 & \frac{-a-1}{2} & 0 & 0 \\ 0 & 0 & -1 & \frac{a-1}{a+1} & \frac{a-1}{a+1} \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & \frac{-a-1}{2} & -1 & -1 \end{bmatrix} \quad (155)$$

These matrices are such that $I^t M_j = M_{j+1}$, with indices taken mod 2.

To recover the arithmetic graph of an orbit, in the sense of the Graph Master Picture Theorem, we take the lattice path in \mathbf{Z}^5 and apply the linear transformation

$$(x_1, x_2, x_3, x_4, y) \rightarrow \left(\frac{x_3}{2}, \frac{x_2 + x_3 + x_4}{2} \right). \quad (156)$$

When $A = p/q$ we can recover the entire arithmetic graph by choosing $\iota > 0$ sufficiently small, say $\iota < 1/q^2$, and setting $V_0 = (\iota, \iota, \iota, \iota, \iota)$.

16 Proof of the Structure Theorem

16.1 The Singular Directions

In this chapter we prove Theorem 15.6. Let \mathcal{H} denote a finite union of n -dimensional linear subspaces of \mathbf{R}^{n+1} . We say that \mathcal{H} is a *complete set* for the map $\widehat{\mathcal{T}}$ if $\widehat{\mathcal{T}}$ is defined on $\widehat{\mathbf{S}} - X$, where X is a finite union of codimension 1 flat tori and each element of X is parallel to some element of \mathcal{H} .

In this section we will produce a complete set \mathcal{H} with $n + 1$ members. This is a first step towards proving Theorem 15.6 because the map $[X_1, X_2, I]$ discussed in Theorem 15.6 have complete sets with $n + 1$ members.

Let Π_k denote the hyperplane given by the equation $x_k = 0$. To keep our notation consistent with the previous chapter, we say that Π_{n+1} is the hyperplane given by $y = 0$. Let $L(F)$ denote the linear part of the affine map F .

Lemma 16.1 *A complete set for \mathcal{T} is given by*

1. $\{\Pi_1, \Pi_{n+1}\}$.
2. $L(\widehat{R}_1)^{-1}(\Pi_{n+1})$.
3. $L(\widehat{R}_1)^{-1}L(\widehat{S}_1)^{-1}(\{\Pi_2, \Pi_{n+1}\})$,
4. $L(\widehat{R}_1)^{-1}L(\widehat{S}_1)^{-1}L(\widehat{R}_2)^{-1}(\Pi_{n+1})$,
5. $L(\widehat{R}_1)^{-1}L(\widehat{S}_1)^{-1}L(\widehat{R}_2)^{-1}L(\widehat{S}_2)^{-1}(\{\Pi_3, \Pi_{n+1}\})$,

and so on.

Proof: In view of Lemma 15.3, the hyperplane Π_{n+1} is a complete set for \widehat{S}_j . In view of Lemma 15.4 (and the remarks after it), the two hyperplanes $\{\Pi_k, \Pi_{n+1}\}$ form a complete set for \widehat{R}_k . If the map $\widehat{\mathcal{T}}$ is not defined on some point p , then one of the compositions

$$F_k = \widehat{R}_k \circ \widehat{S}_{k-1} \circ \cdots \circ \widehat{S}_1 \circ \widehat{R}_1, \quad G_k = \widehat{S}_k \circ \widehat{R}_k \circ \cdots \circ \widehat{S}_1 \circ \widehat{R}_1 \quad (157)$$

is undefined at p but all shorter compositions are defined. But then either $F_k(p)$ lies in the boundary of the invariant domain for \widehat{R}_k or $G_k(p)$ lies in the boundary of the invariant domain for \widehat{S}_k . But then p lies in a hypersurface parallel to one of the hyperplanes on our list. ♠

Lemma 16.2 *A complete list for \mathcal{T} is given by H_1, \dots, H_{n+1} , where $H_1 = \Pi_1$ and $H_{n+1} = \Pi_{n+1}$ and*

$$H_{k+1} = L(\widehat{R}_1)^{-1} \circ \dots \circ L(\widehat{S}_k)^{-1}(\Pi_{k+1}), \quad k = 1, \dots, n-1. \quad (158)$$

Proof: Note that there are about $3n$ hyperplanes listed in Lemma 16.1 whereas we are claiming that $n+1$ hyperplanes suffices. The idea here is just to eliminate redundancies. First, we have

$$L(\widehat{R}_k)^{-1}(\Pi_{n+1}) = \Pi_k. \quad (159)$$

Therefore, each hyperplane listed on line $2k$ of Lemma 16.1 is contained in one of the hyperplanes listed on line $2k-1$.

Second, we have

$$L(S_k)^{-1}(\Pi_{n+1}) = \Pi_{n+1}, \quad L(\widehat{R}_k)^{-1}(\Pi_k) = \Pi_{n+1}. \quad (160)$$

Therefore, the second hyperplane listed on line $2k+1$ of Lemma 16.1 is contained first hyperplane listed on line $2k-1$. For instance, taking $k=2$, we have

$$\begin{aligned} L(\widehat{R}_1)^{-1}L(\widehat{S}_1)^{-1}L(\widehat{R}_2)^{-1}L(\widehat{S}_2)^{-1}(\Pi_{n+1}) = \\ L(\widehat{R}_1)^{-1}L(\widehat{S}_1)^{-1}L(\widehat{R}_2)^{-1}(\Pi_{n+1}) = L(\widehat{R}_1)^{-1}L(\widehat{S}_1)^{-1}(\Pi_2). \end{aligned}$$

Upon eliminating all the redundancies, we get the advertised list. ♠

Let e_k denote the k th standard basis vector in \mathbf{R}^{n+1} . Let H_k^\perp denote the normal to H_k .

Lemma 16.3 *The matrix whose rows are $H_1^\perp, \dots, H_{n+1}^\perp$ has determinant 1.*

Proof: Let $M_k = L(\widehat{R}_1) \circ \dots \circ L(\widehat{S}_{k-1})$. We have $H_n^\perp = (0, \dots, 0, 1)$ and

$$H_k^\perp = (M_k^{-1})^t(e_k), \quad k = 1, \dots, n. \quad (161)$$

The maps $L(\widehat{R}_j)$ and $L(\widehat{R}_k)$ act trivially on e_{j+1}, \dots, e_n . Hence M_k acts trivially on e_{k+1}, \dots, e_n . Hence, rows k, \dots, n of the inverse transpose matrix $(M_k^{-1})^t$ coincide with the rows of the identity matrix. Hence

$$H_k^\perp = (*, \dots, *, 1, 0, \dots, 0, *), \quad k = 1, \dots, n. \quad (162)$$

The 1 appears in the k th slot and $(*)$ indicates an entry that we don't explicitly know. The lemma is immediate from this structure. ♠

16.2 The First Parallelotope

Let $X_1 \subset \mathbf{R}^{n+1}$ be the parallelotope consisting of vectors V such that

$$H_i^\perp \cdot V \in [-1/2, 1/2] \quad (163)$$

for all i .

Lemma 16.4 X_1 is a fundamental domain for \mathbf{Z}^{n+1} .

Proof: In view of Lemma 16.3, the set X_1 is a unit volume parallelotope. Let M be the matrix with rows $H_1^\perp, \dots, H_{n+1}^\perp$. From the proof of Lemma 16.3 we have

$$M = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & * \\ * & 1 & 0 & \cdots & 0 & * \\ * & * & 1 & \cdots & 0 & * \\ \cdots & & & & & \\ * & * & * & \cdots & 0 & 1 \end{bmatrix} \quad (164)$$

X_1 consists of those vectors $V \in \mathbf{R}^{n+1}$ such that $MV \in [-1/2, 1/2]^{n+1}$.

Since X_1 has unit volume, it suffices to show that the interior of X_1 does not intersect some integer translate of X_1 . This happens if and only if there is some integer vector $V \in \mathbf{Z}^{n+1}$ such that $MV \in (0, 1)^{n+1}$. This is clearly impossible given the form of M . ♠

Remark: Given the form of the matrix in Equation 164, we can also say that X_1 is the polytope bounded by the hyperplanes $H_k \pm 1/2e_k$.

Let $q = (q_1, \dots, q_n)$ and $r = (\dots)$ and $s = (\dots)$ be the invariants for \mathcal{T} . Let $\pi : \mathbf{R}^{n+1} \rightarrow \widehat{\mathbf{S}}$ be projection. Let X_1^o be the interior of X_1 . Let I be the affine map which fixes the vector $q/2$ and whose linear part coincides with the linear part of $\widehat{\mathcal{T}}$.

Lemma 16.5 the map $\widehat{\mathcal{T}}$ is entirely defined on $\pi(X_1^o + q/2)$ and $\widehat{\mathcal{T}} = \pi \circ I \circ \pi^{-1}$ on $\pi(X_1^o + q/2)$ provided that π^{-1} is taken to have its range in $X_1^o + q/2$.

Proof: We will give the proof in case $q = (0, \dots, 0)$. In this case, I is simply the linear part of $\widehat{\mathcal{T}}$. The general case has essentially the same proof, and differs only in that we apply suitable translations to the basic objects.

Let A_k denote the open slab bounded by the hyperplanes $x_k = \pm 1/2$. Let B_k denote the open slab bounded by the parallel hyperplanes $H_k \pm \frac{1}{2}e_k$. By construction $X_1 = \bigcap H_k$. Also by construction,

$$L(\widehat{R}_1)^{-1} \circ \cdots \circ L(\widehat{S}_{k-1})^{-1}(A_k) = B_k. \quad (165)$$

Let ρ_k be the restriction of $L(\widehat{R}_k)$ to $A_k \cap A_{n+1}$. Likewise, let σ_k be the restriction of $L(\widehat{S}_k)$ to A_{n+1} . Given the description of the invariant domains for \widehat{R}_k and \widehat{S}_k in §15.3, we have

$$\widehat{R}_k = \pi \circ \rho_k \circ \pi^{-1}, \quad \widehat{S}_k = \pi \circ \sigma_k \circ \pi^{-1}. \quad (166)$$

The right hand side is independent of the lift, as long as the range of π^{-1} is taken to be $A_k \cap A_{n+1}$ or A_{n+1} respectively.

Choose any point $p \in \pi(X_1^o)$. Let q_1 be the unique point in X_1^o such that $\pi(q_1) = p$. By construction $q_1 \in B_1 \cap \cdots \cap B_{n+1}$. But $A_1 = B_1$ and $A_{n+1} = B_{n+1}$. Hence $q_1 \in A_1 \cap A_{n+1}$. Since $q_1 \in A_1 \cap A_{n+1}$, the map ρ_1 is defined on q_1 . Since ρ_1 preserves $A_1 \cap A_{n+1}$, we have $\rho_1(q_1) \in A_1 \cap A_{n+1}$. In particular $\rho_1(q_1) \in A_{n+1}$, and so σ_1 is defined on $\rho_1(q_1)$. Equation 166 now gives us

$$q_2 = \sigma_1 \circ \rho_1(q_1) \in A_2 \cap A_{n+1}, \quad \pi(q_2) = \widehat{S}_1 \circ \widehat{R}_1(p). \quad (167)$$

Repeating the same argument with q_2 in place of q_1 , we see that ρ_2 is defined on q_2 and σ_2 is defined on $\rho_2(q_2)$ and

$$q_3 = \sigma_2 \circ \rho_2(q_2) \in A_3 \cap A_{n+1}, \quad \pi(q_3) = \widehat{S}_2 \circ \widehat{R}_2 \circ \widehat{S}_1 \circ \widehat{R}_1(p). \quad (168)$$

Continuing in this way, we produce points q_4, \dots, q_n such that

- $q_k \in A_k \cap A_{n+1}$.
- $\sigma_k \circ \rho_k$ is defined on q_k .
- $q_{k+1} = \sigma_k \circ \rho_k(q_k)$.
- $\pi \circ q_{k+1} = \widehat{S}_k \circ \widehat{R}_k(\pi(q_k))$.

In particular, $\widehat{\mathcal{T}}$ is defined on p and

$$\widehat{\mathcal{T}}(p) = \pi(q_n) = \pi \circ \sigma_n \circ \cdots \circ \rho_1 \circ \pi^{-1}(q_1) = I \circ \pi^{-1}(p). \quad (169)$$

Hence $\widehat{\mathcal{T}}$ is completely defined on $\pi(X_1^o)$ and $\widehat{\mathcal{T}} = \pi \circ I \circ \pi^{-1}$ on $\pi(X_1^o)$. ♠

16.3 The Second Parallelotope

Let $X_2 = I(X_1)$.

Lemma 16.6 X_2 is a fundamental domain for \mathbf{Z}^{n+1} .

Proof: Again, we consider the case when $q = (0, \dots, 0)$ for ease of exposition. The linear parts of \widehat{R}_k and \widehat{S}_k are orientation preserving and volume preserving maps. We also know that X_1 is a unit volume parallelotope and a fundamental domain for \mathbf{Z}^{n+1} . Since I is volume preserving, X_2 is also a unit volume parallelotope.

The map $\widehat{\mathcal{T}}$ is invertible. In particular, the restriction of $\widehat{\mathcal{T}}$ to X_1^o is injective. But this map equals $\pi \circ I \circ \pi^{-1}$. Hence $\pi : X_2 \rightarrow \widehat{\mathbf{S}}$ is also injective. This fact, together with the fact that X_2 has unit volume, shows that X_2 is in fact a fundamental domain. ♠

When $q = (0, \dots, 0)$, Lemma 16.5 tells us that $\widehat{\mathcal{T}} = [X_1, X_2, I]$. In general, let $X'_j = X_j + q/2$ and let I' be the affine map which fixes $q/2$ and whose linear part is I . Lemma 16.5 tells us that $\widehat{\mathcal{T}} = \pi \circ I' \circ \pi^{-1}$ on the interior of $\pi(X'_1)$. But $[X'_1, X'_2, I']$ is conjugate to $[X_1, X_2, I]$.

More precisely, let $\tau : \widehat{\mathbf{S}} \rightarrow \widehat{\mathbf{S}}$ be translation by $q/2$. Then

$$\tau \circ [X_1, X_2, I] \circ \tau^{-1} = [X'_1, X'_2, I'].$$

It is convenient to define the new map

$$\Psi_q = \tau^{-1} \circ \Psi = \Psi - q/2. \quad (170)$$

A short calculation tells us that

$$\Psi_q \circ \mathcal{T} = [X_1, X_2, I] \circ \Psi_q. \quad (171)$$

From this alternate point of view, the compactified system $[X_1, X_2, I]$ is independent of the q parameters. What changes with the q parameters is the map Ψ_q .

It is worth remarking that the map $p \rightarrow -p$ gives an involution on $\widehat{\mathbf{S}}$ having 2^{n+1} fixed points. Of these fixed points, 2^n are distinguished by the property that the last coordinate is $[0]$. The map Ψ_q maps the origin to one of these distinguished fixed points.

16.4 The Fixed Point Set

It only remains to prove that I pointwise fixes a codimension 2 subspace. It suffices to consider the case $q = (0, \dots, 0)$. We call a QTC *good* if Ψ is dense in $\widehat{\mathbf{S}}$ and the linear part of \mathcal{T} has 2 unequal eigenvalues.

Lemma 16.7 *The set of good QTCs is dense in \mathbf{R}^{2n} .*

Proof: The set of QTCs satisfying the second condition has full measure. All we need here is that there are no rational relations amongst the numbers $r_n/r_1, \dots, r_n/r_{n-1}, 1$. The set of QTCs satisfying the first condition has the form $F^{-1}(0)$ where F is a polynomial function. The point is that the trace of the linear part of \mathcal{T} is a polynomial in the variables, and we just need to avoid the trace value 2. If we can show that the set of QTCs satisfying the first condition is nonempty, then this set is open dense. Finally, the intersection of an open dense set with a full measure set is dense.

In case n is not divisible by 4, the QTC with parameters $r = (1, \dots, 1)$ and $s = (0, \dots, 0)$ has the desired properties. When n is divisible by 4, the QTC with parameters $r = (1, \dots, 1, 2)$ and $s = (0, \dots, 0)$ has the desired properties.

♠

The linear map I depends continuously on the QTC parameters, so it suffices to consider the case when \mathcal{T} is good.

Lemma 16.8 *Let A be an area preserving affine map of \mathbf{R}^2 with unequal eigenvalues. Let $\epsilon > 0$ be given. Then there is some $\delta > 0$ such that the bound $\|A(p) - p\| < \delta$ implies that A has a fixed point within ϵ of p .*

Proof: This is a standard result. It suffices to consider the case when A is linear. But, as is well known, the only points which such a map almost fixes are near the origin. ♠

Recall that we have the triple (X_1, X_2, I) associated to \mathcal{T} . The map I certainly fixes the origin. There is some small ball $X'_1 \subset X_1$ centered at the origin such that $I(X'_1) \subset X_1$.

Say that a *net* of \mathbf{S} is a subset of points such that every point of \mathbf{S} is within some N of a point in the subset. Let $\Theta \subset \mathbf{S}$ denote the set of points $(x, y) \in \mathbf{S}$ such that $\Psi(x, y) \in X'_1$ and \mathcal{T} fixes (x, y) .

Lemma 16.9 Θ is a net in \mathbf{S} .

Proof: Associated to the quarter turn maps R_1, \dots, R_n there are rectangle tilings $\mathcal{R}_1, \dots, \mathcal{R}_n$. For any $\epsilon > 0$ we can find a net $X \subset \mathbf{S}$ with the following property. Each $(x, y) \in X$ is within ϵ of a center of a rectangle from each tiling. If ϵ is small enough then \mathcal{T} is defined on a ball of radius $\epsilon_0 > 0$ about (x, y) and moves (x, y) no more than ϵ . Here ϵ_0 is a universal constant that does not tend to 0 with ϵ . For such points, the distance from $\Psi(x, y)$ to $(0, \dots, 0)$ tends to 0.

If we choose ϵ small enough then there is some $\epsilon_1 > 0$ such that \mathcal{T} fixes a point (x', y') within ϵ_1 of each $(x, y) \in X$. This follows from Lemma 16.8 applied to the restriction of \mathcal{T} to the ϵ_0 -ball about (x, y) . Here ϵ_1 tends to 0 with ϵ . If ϵ_1 is sufficiently small then $\Psi(x', y') \in X'_1$. The set of all such (x, y') forms the desired net. ♠

Let $\Pi' \subset X'_1$ denote the intersection of the +1 eigenspace of I with X'_1 . By construction I fixes a point in X'_1 if and only if that point lies in Π' . Therefore, $\Psi(\Theta) \subset \Pi'$. We will suppose that $\dim(\Pi) < n - 1$ and derive a contradiction.

Let $\Pi \subset \mathbf{R}^{n+1}$ denote the linear subspace spanned by tangent vectors to Π' the two vectors $d\Psi(1, 0)$ and $d\Psi(0, 1)$. Then Π is a proper linear subspace of \mathbf{R}^{n+1} and has measure 0. By the ergodic theorem, the set

$$X = \Psi^{-1}(\Pi \cap \widehat{\mathbf{S}}) \tag{172}$$

has density 0 in \mathbf{S} . On the other hand, since Ψ is locally affine, X contains the ϵ neighborhood of Θ for some $\epsilon > 0$. Since Θ is a net, and $\epsilon > 0$, we see that X has positive density in \mathbf{S} . This is a contradiction. We conclude that $\dim(\Pi') \geq n - 2$. This is what we wanted to prove.

Remark: In case \mathcal{T} comes a QTC derived from a pinwheel map, the linear part of \mathcal{T} is the reflection through the origin. In this case, I acts on the 2-dimensional set $d\Psi(\mathbf{R}^2)$ as reflection through the origin. This shows that I has two -1 eigenvalues. Since we already know that I has $n - 1$ eigenvalues of value 1, we see that the complete list of eigenvalues of I must be $-1, -1, 1, \dots, 1$.

Part IV

The Quasi-Isomorphism Theorem

17 The Proof in Broad Strokes

17.1 Pixellated Squares

The purpose of this part of the monograph is to prove the Quasi-Isomorphism Theorem. We fix some even rational parameter $A = p/q$ for the entire chapter. All the definitions are made with respect to this parameter. Also, to save words, *square* will always denote a unit integer square. Finally, by *arithmetic graph*, we mean the image under the canonical affine transformation.

Plaid Nontriviality: We say that a square Σ is *plaid nontrivial* if the plaid tile at Σ is nontrivial. Otherwise we call Σ *plaid trivial*. When Σ is plaid nontrivial, we define the *plaid edge set* of Σ to be the two edges of Σ crossed by the plaid polygon that enters Σ .

Grid Fullness: We say that Σ is *grid full* if Σ contains a point of the graph grid. Otherwise we call Σ *grid empty*. When Σ is grid full, there is a unique point of the graph grid contained in Σ , and this point lies in the interior of Σ . See the Grid Geometry Lemma.

Graph Nontriviality: This definition only applies when Σ is grid full. We say that Σ is *graph trivial* if the point σ of the graph grid contained in Σ is isolated in the arithmetic graph. Otherwise we call Σ *graph nontrivial*. When Σ is graph nontrivial, we call the two edges of the arithmetic graph incident to σ the *graph edges associated to Σ* .

Pixellation: This definition only applies to squares Σ which are graph full. We call Σ *pixellated* if the following is true:

- Σ is graph trivial if and only if Σ is plaid trivial.
- If Σ is plaid nontrivial then the graph edges associated to Σ cross Σ in the interiors of the edges in the plaid edge set of Σ .

When Σ is pixellated, the plaid model at Σ determines the local picture of the graph model in Σ in the cleanest possible way.

We can see that all the grid full squares in Figure 1 are pixellated. Indeed, for the parameter $3/8$ and many others, every grid full square is pixellated. However, this perfect situation fails for some parameters.

17.2 Bad Squares

Bad Squares We keep the notation and terminology from the previous section. We call the square Σ *bad* if Σ is graph full but not pixellated. It turns out that this only happens when Σ is both plaid nontrivial and graph nontrivial. (See the Pixellation Theorem below.)

Offending Edges: We say that an *offending edge* is a graph edge associated to Σ which does not cross the boundary of Σ in the interior of one of the edges of the edge set.

Unused Sides: We say that an *unused side* is an edge in the plaid edge set of Σ which is not crossed by a graph edge associated to Σ . The existence of an unused side implies the existence of an offending edge, and *vice versa*.

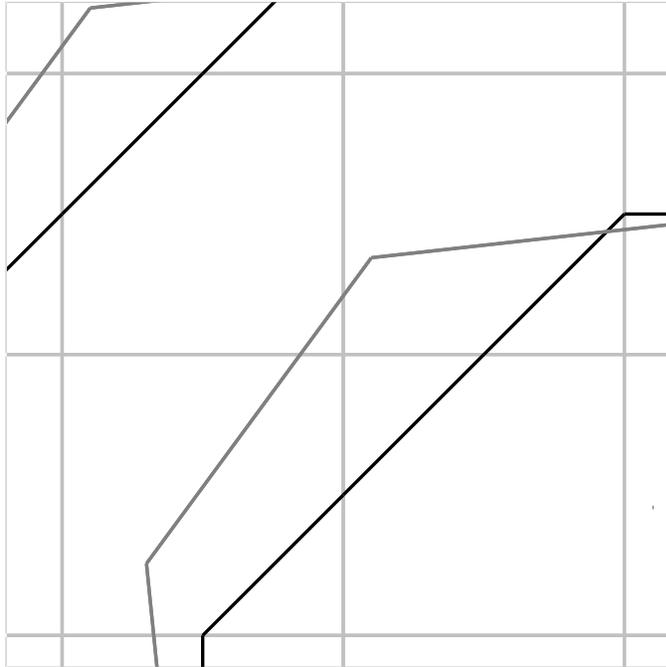


Figure 17.1: Some bad squares for the parameter $p/q = 4/5$.

Figure 17.1 shows a closeup of the picture when $p/q = 4/5$. The plaid polygons are in black and the arithmetic graph polygons are in grey. The top right and bottom left squares are bad. The top left and bottom right squares are grid empty.

Catches: Figure 17.2 shows a picture of two possible patterns of squares which involve a bad square, an offending edge and an unused side. In both cases, the bad square is meant to be the top right one. The plaid segment in the top right square either connects the south edge to the north edge or to the east edge. The shaded squares are meant to be grid empty. The offending edge is drawn as a curved grey segment. We mean to consider now just these patterns, but also the ones which arise by applying a symmetry of the square grid to these. In other words, we don't want to fix the orientation of the pattern in the plane.

In both cases, the portion of the plaid model in the bottom two squares makes a straight diagonal segment, and the sign of the slope of the offending edge is the same as the sign of the portion of the diagonal segment. In these pictures, we say that the unused side and the offending edge are *associated*. We call these two patterns (and their isometric images) *catches* for the offending edge.

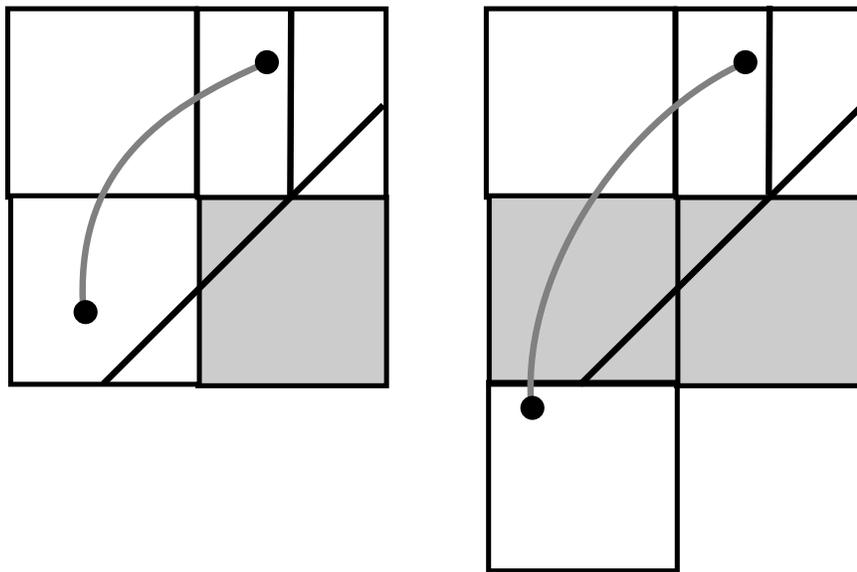


Figure 17.2: The two catches for an offending edge in a bad square.

The Pixellation Theorem, stated below, says that the pictures in Figure 17.2 and their rotated/reflected images are always present when there is an unused side and an offending edge. We defer the statement of the Pixellation Theorem for a while, because we want to bundle some other minor results into it.

17.3 Errand Edges and Double Crossings

Errand Edges: Say that an *errand edge* is an edge of the arithmetic graph with sits with respect to the plaid model as in Figure 17.3. The grey arc is the arithmetic graph edge. We have drawn it curved because here and below a curved picture will look nicer. Of course, we are representing a straight line segment. The numbers 1 and 2 denote squares Σ_1 and Σ_2 . The edge $\Sigma_1 \cap \Sigma_2$ is in the plaid edge set of Σ_1 (and of Σ_2) and hence the arithmetic graph edge is not an offending edge with respect to Σ_1 . However, it rises up at least one unit above the top of Σ_1 and Σ_2 while the plaid polygon in Σ_2 either goes straight across Σ_2 or else moves downward. As usual, we mean to consider all possible orientations of these configurations.

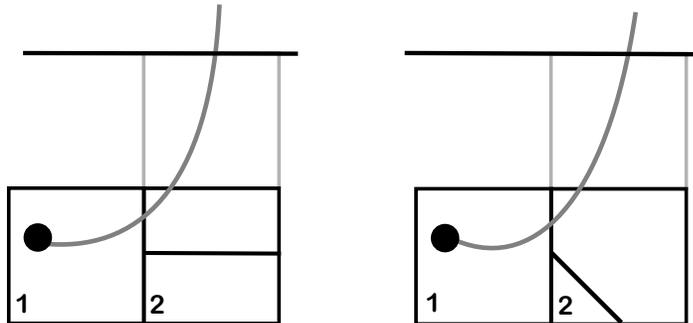


Figure 17.3: errand edges

Double Crossings: We say that a *double crossing* is a union of two disjoint distinguished edges e_1 and e_2 which have endpoints v_1 and v_2 in adjacent unit integer squares Σ_1 and Σ_2 and both cross $\Sigma_1 \cap \Sigma_2$ at interior points. The two squares may either be stacked on top of each other, as in Figure 17.4, or stacked side by side.

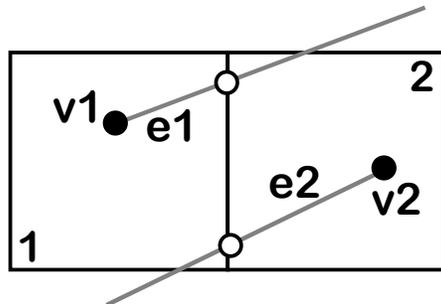


Figure 17.4: A double crossing

17.4 The Pixellation Theorem

Here is our main result.

Theorem 17.1 (Pixellation) *The following is true for any even rational parameter.*

1. *There are no double crosses in the arithmetic graph.*
2. *A square is plaid nontrivial if and only if it is grid nontrivial.*
3. *There are no errant edges in the arithmetic graph.*
4. *When a square is graph nontrivial, the two associated graph edges must cross distinct sides of the square.*
5. *In a bad square, there is a bijection between unused sides and offending edges, and each matched pair of objects is involved in a catch.*

The Pixellation Theorem says that the vast majority of grid full squares are pixellated, and it precisely characterizes the local picture around the ones which are not. To deduce the Quasi-Isomorphism Theorem from the Pixellation Theorem, we need to see that the various local pictures implied by the Pixellation Theorem piece together correctly. The next section contains the result we need, the Bound Chain Lemma. Here are some remarks on the wording in the Pixellation Theorem.

Statement 5 has a complicated phrasing which we want to explain. Logically, the existence of an unused side implies the existence of an offending edge, but we want to make sure that there is an offending edge that specifically is associated to the unused side of interest to us. Likewise, the existence of an offending edge implies the existence of an unused side, but we want to make sure that there is an unused side that specifically is associated to the offending edge of interest to us.

Statements 1 and 3 will not be used directly in our deduction of the Quasi-Isomorphism Theorem. However, they will be used, in the next chapter, for the deduction of the Bound Chain Lemma.

17.5 The Bound Chain Lemma

The main weakness in the Pixellation Theorem is that it only deals with grid full squares. Here we deal with the grid empty squares.

We say that a finite union of squares $\Sigma_1, \dots, \Sigma_m$ is *linked* if

- Σ_1 and Σ_m are grid fill, and the remaining squares are grid empty.
- $\Sigma_k \cap \Sigma_{k+1}$ is an edge, for each $k = 1, \dots, (m - 1)$.
- A single plaid polygon intersects Σ_k for all k .

To avoid trivialities, we take $m \geq 2$. It follows from Statements 3 and 4 of the Grid Geometry Lemma that a linked chain has length at most 4.

We say that our linked sequence is *bound* if a single edge in the arithmetic graph joins the graph grid point in Σ_1 to the graph grid point in Σ_m .

Figure 17.5 shows a linked and bound chain of length 4. We have drawn the arithmetic graph edge as curved, to get a nicer picture.

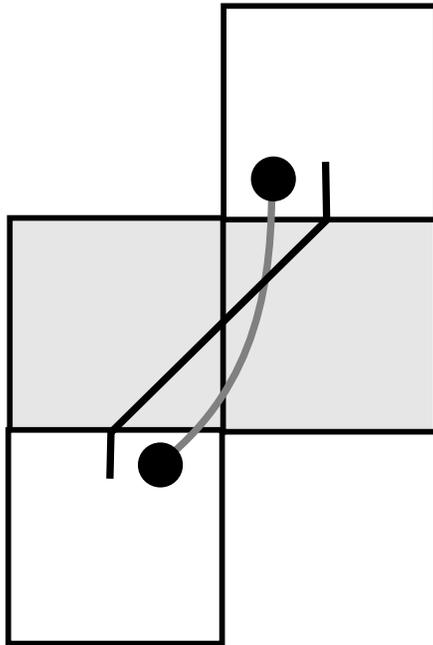


Figure 17.5: A Linked and Bound Chain of Length 4

In the next chapter, we will prove the following result, essentially using the Pixellation Theorem and a case-by-case analysis.

Lemma 17.2 (Bound Chain) *Every linked chain is bound.*

17.6 Proof of the Quasi-Isomorphism Theorem

The Quasi-Isomorphism Theorem is a fairly immediate consequence of the Pixellation Theorem and the Bound Chain Lemma. Here we produce an explicit homeomorphism between the union of polygons in the plaid model and the union of polygons in the arithmetic graph. The homeomorphism moves no point more than 2 units. The homeomorphism is defined in several pieces, depending on the type of square we have.

Let Σ be a pixellated square. Let v be the graph grid point in Σ and let e_1 and e_2 be the two edges incident to v . Let v' be the midpoint of the plaid segment contained in Σ and let e'_1 and e'_2 be the halves of the plaid segment on either side of v' . We map $e_k \cap \Sigma$ linearly to e'_k for $k = 1, 2$. We choose the labeling so that e_k and e'_k both intersect the same side of Σ . Figure 17.6 shows this simple map in action.

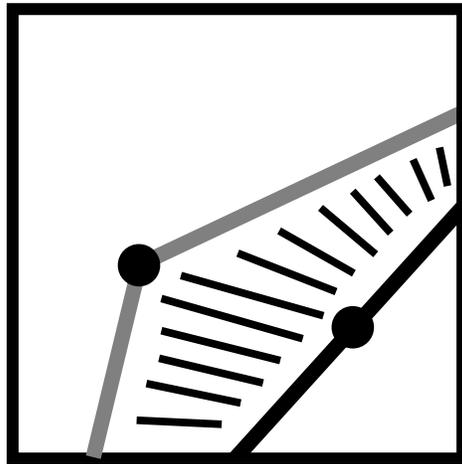


Figure 17.6: The Map on Pixellated Squares

We use the same notation for discussing bad squares. If Σ is a bad square and e_1 is not an offending edge of Σ then we do the same map as for pixellated squares. If e_1 is an offending edge, we map e_1 to the union of squares in the catch as shown in Figure 17.7 for each of the two kinds of catches. The white dots in Figure 17.7 are just extra guides for the map. On the right hand side of Figure 17.7, we have drawn one particular way that the plaid segment could look. The black dot the bottom square is meant to be the midpoint of the plaid segment, whatever it looks like.

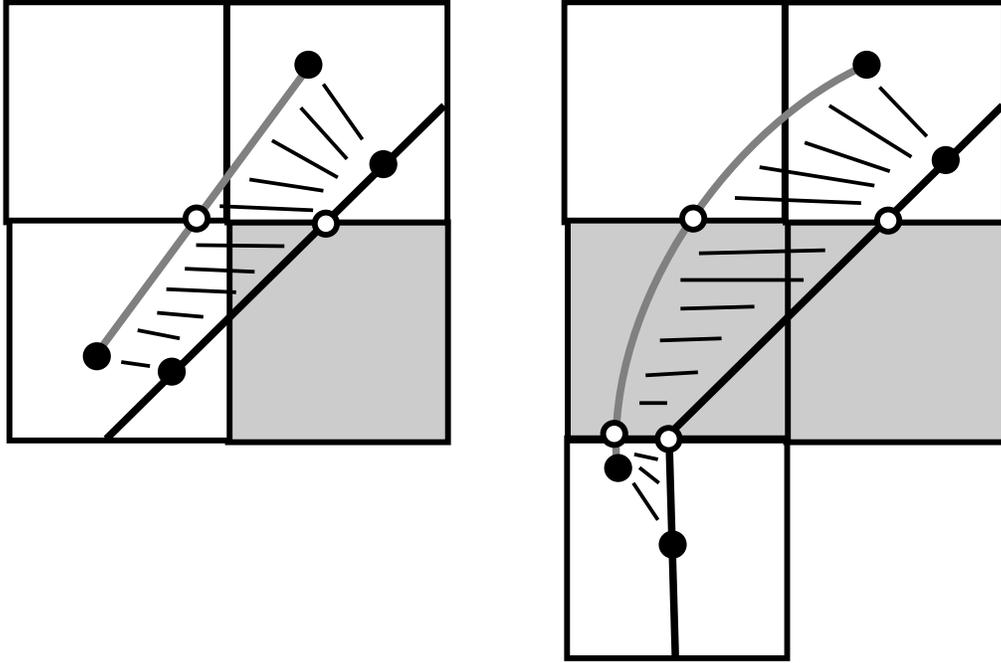


Figure 17.7: The Map on Bad Squares

Say that a *clean linked chain* is a chain $\Sigma_1, \dots, \Sigma_m$ such that the arithmetic graph edge e connecting v_1 to v_m crosses $\Sigma_1 \cap \Sigma_2$ and $\Sigma_{m-1} \cap \Sigma_m$. We say that the *graph core* associated to the clean linked chain is the segment

$$\hat{e} = e - (\Sigma_1 \cap \Sigma_m). \quad (173)$$

The segment \hat{e} has endpoints on $\Sigma_1 \cap \Sigma_2$ and $\Sigma_{m-1} \cap \Sigma_m$. Corresponding to the \hat{e} is the portion of the plaid model connecting these same two edges. We call this the *plaid core* associated the clean linked chain.

Our map is defined everywhere on the arithmetic graph polygons except on the graph cores. These comprise a disjoint set of segments. The image of our map so far is exactly the complement of the plaid cores. We finish the proof by mapping the graph cores to the plaid cores according to the scheme in Figure 17.8.

Figure 17.8 doesn't show every possibility, but these examples should be sufficient to show what we do in every case. In every case, the Pixellation Theorem and the Bound Chain Theorem imply that the map is well defined and only moves points by at most 2 units. We call our map Θ . By construction, Θ maps each arithmetic graph polygon homeomorphically to some plaid polygon.

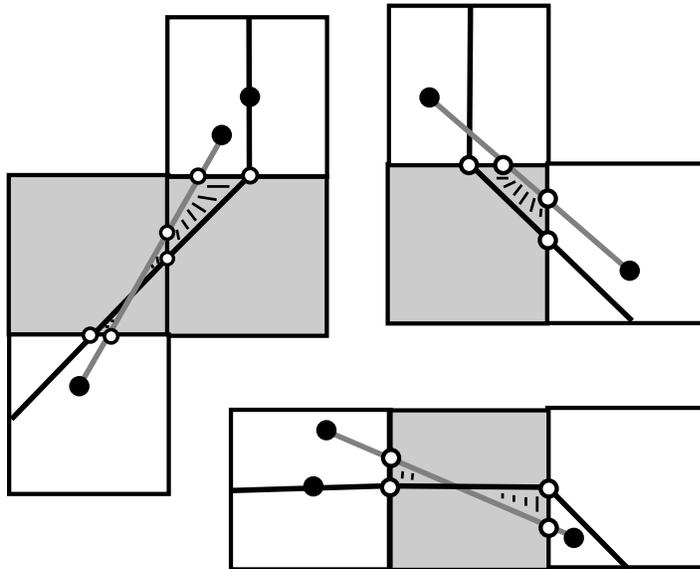


Figure 17.8: The Map on Bad Squares

Lemma 17.3 Θ is injective.

Proof: Suppose γ_1 and γ_2 are graph polygons which both map to the plaid polygon π . Suppose π never enters a square with a non-offending edge. Then, according to the Pixellation Lemma, π travels in a straight diagonal line. Since π is closed, this situation is impossible. Hence, π must enter at least one square Σ which has at least one non-offending edge e . The only way γ_1 and γ_2 are both mapped on to $e \cap \Sigma$ is if both these polygons contain the grid point in Σ . But then $\gamma_1 = \gamma_2$. ♠

Lemma 17.4 Θ is surjective.

Proof: Each plaid polygon π has length at least 4. Hence, by Statements 3 and 4 of the Grid Geometry Lemma the polygon π intersects at least one grid full square Σ . But then Φ maps the arithmetic graph polygon which contains the graph grid point in Σ to π . ♠

In short, Θ is a bijection between the components and a homeomorphism on each one, and Θ moves points by at most 2 units. This completes the proof of the Quasi-Isomorphism Theorem.

18 Proof of the Bound Chain Lemma

18.1 Length Two Chains

Let Σ_1, Σ_2 be a linked chain of length 2. This means that both Σ_1 and Σ_2 are plaid nontrivial and grid full, and a single plaid polygon runs through both. By the Pixellation Theorem, Σ_1 and Σ_2 are graph full as well.

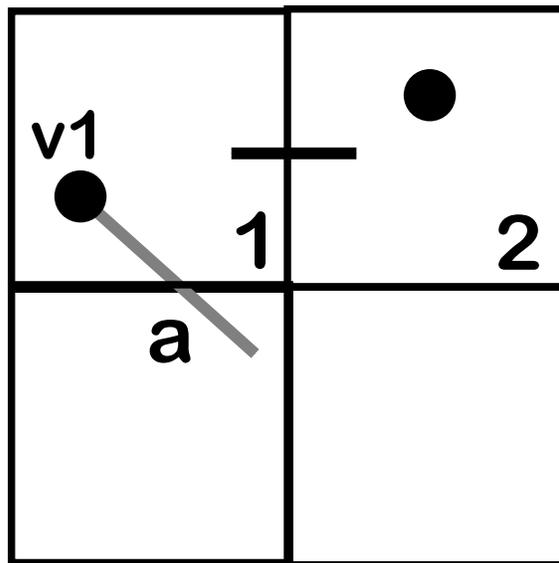


Figure 18.1: Σ_1 and Σ_2 , and an uncaught offending edge.

Let v_j be the graph grid point in Σ_j . Suppose that no graph edge incident to v_1 crosses $\Sigma_1 \cap \Sigma_2$. Then there must be an offending edge a incident to v_1 and associated to $\Sigma_1 \cap \Sigma_2$. This edge must cross either the bottom or the top of Σ_1 , and we have shown it crossing the bottom. But the picture contradicts Statement 5 of the Pixellation Theorem, because the catch for a would involve Σ_2 , and Σ_2 would be grid empty in the catch.

The argument in the preceding paragraph shows that there is some graph edge e_1 incident to v_1 which crosses $\Sigma_1 \cap \Sigma_2$. Likewise, there is some graph edge e_2 incident to $v_2 \in \Sigma_2$ which crosses $\Sigma_1 \cap \Sigma_2$. If e_1 and e_2 are disjoint, then we have a bad configuration. The Bad Configuration Lemma rules this out. Since the arithmetic graph is embedded, one of our two edges connects v_1 to v_2 , and in fact these two edges must coincide. (Otherwise we contradict the Pixellation Theorem.) In short, Σ_1, Σ_2 is a bound chain.

18.2 Length Three Chains: Case A

Here we consider a length 3 chain in which the three squares are the same horizontal row.

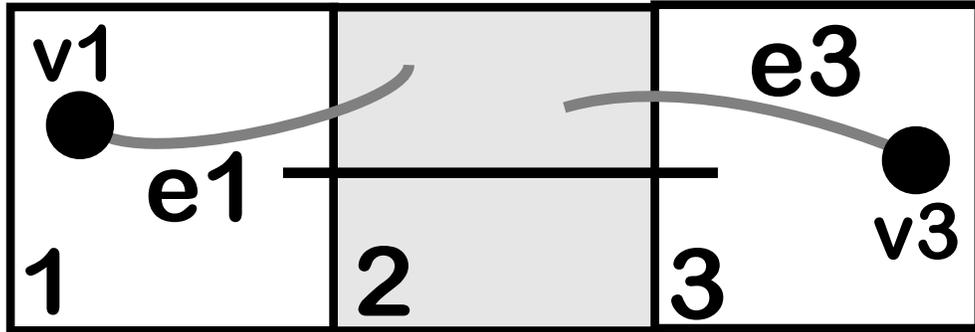


Figure 18.2 A horizontal chain

Lemma 18.1 *There is an edge of the arithmetic graph connecting v_1 and v_3 .*

Because the plaid polygon in Σ_2 is not a diagonal segment, there is no way to involve Σ_2 in a catch for an offending edge incident to v_1 and associated to $\Sigma_1 \cap \Sigma_2$. Hence, $\Sigma_1 \cap \Sigma_2$ is not an unused edge of Σ_1 . Hence there is some graph edge e_1 incident to v_1 and crossing $\Sigma_1 \cap \Sigma_2$. The same argument shows that there is an arithmetic graph edge e_3 associated to Σ_3 which crosses $\Sigma_2 \cap \Sigma_3$.

If the other endpoint e_1 of e_1 lies in Σ_3 then this other endpoint must be v_3 , because there is at most one graph grid point in Σ_3 . So, to finish the proof, we just have to rule out the other possible squares which could potentially contain the endpoint of e_1 .

The edge e_1 has length less than 2 because $\|dT\|_2 < \sqrt{2}$ and e_1 is the image of a vector in \mathbf{Z}^2 of length at most $\sqrt{2}$. This bound on the length cuts down on the possibilities where e_1 could end up. For one thing, e_1 cannot end up in any square to the right of Σ_3 . Also, e_1 cannot cross the line one unit above the top of Σ_2 , by Statement 3 of the Pixellation Theorem: no errant edges. Hence, the other endpoint of e_1 lies in the row of squares above Σ_2 . We will consider the 3 cases when e_1 ends in a square in the row above Σ_3 . The “below” case has the same treatment.

Case 1: Suppose first e_1 ends in Σ_4 , as shown on the left in Figure 18.3. In this case, e_1 connects v_1 with the graph grid point v_4 in Σ_4 . The plaid edge set of Σ_4 cannot contain $\Sigma_2 \cap \Sigma_4$ because the $\Sigma_2 \cap \Sigma_4$ is not in the plaid edge set of Σ_2 . Hence e_1 is an offending edge for Σ_4 . But then, by the Pixellation Theorem, the portion of the plaid model inside Σ_1 must look as drawn. This is impossible, because it forces the plaid edge set of Σ_1 to have 3 edges in it.

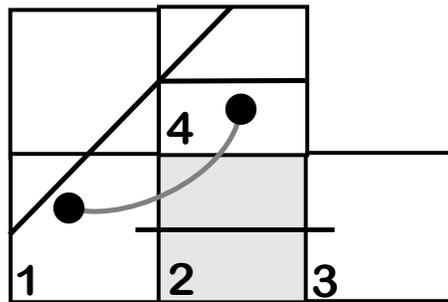


Figure 18.3: Case 1.

Case 2: Suppose e_1 crosses $\Sigma_3 \cap \Sigma_5$ and ends in Σ_5 . Note that e_1 cannot be an offending edge for Σ_5 because Σ_2 does not have the right form to be part of a catch for e_1 . Hence $\Sigma_3 \cap \Sigma_5$ is in the plaid edge set for Σ_5 . But then the result for length 2 chains says that some graph edge joins v_3 to v_5 . But then both graph edges incident to v_5 cross the same edge of Σ_5 . This contradicts Statement 4 of the Pixellation Lemma.

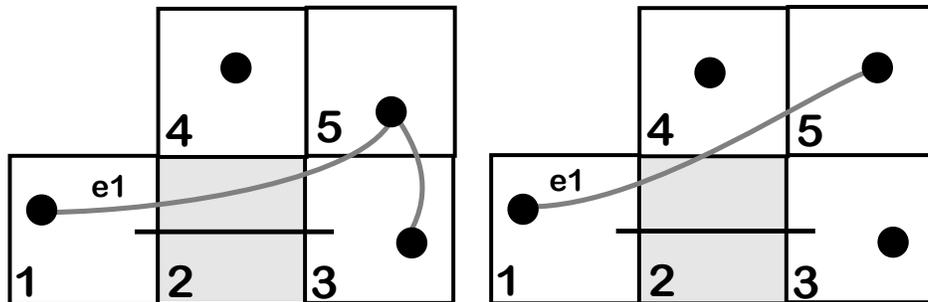


Figure 18.4: Cases 2 and 3

Case 3: Suppose that $w_1 \in \Sigma_5$ and e_1 crosses $\Sigma_4 \cap \Sigma_5$. By Statement 4 of the Grid Geometry Lemma, the square Σ_4 is grid full, just like Σ_3 is. The same argument as in Case 2, with Σ_4 replacing Σ_3 , takes care of this case. ♠

18.3 Length Three Chains: Case B

Here we consider a length 3 chain in which the three squares are the same vertical column. Our proof refers to Figure 18.5

This situation is not quite the same as in Case A because we can have 2 horizontally consecutive grid empty squares. If neither Σ_4 nor Σ_5 , shown in Figure 18.5, is grid empty, then the same argument as in Case A applies here. We just have to worry about the case when one of Σ_4 or Σ_5 is grid empty. We will consider the case when Σ_5 is grid empty. The other case has the same treatment.

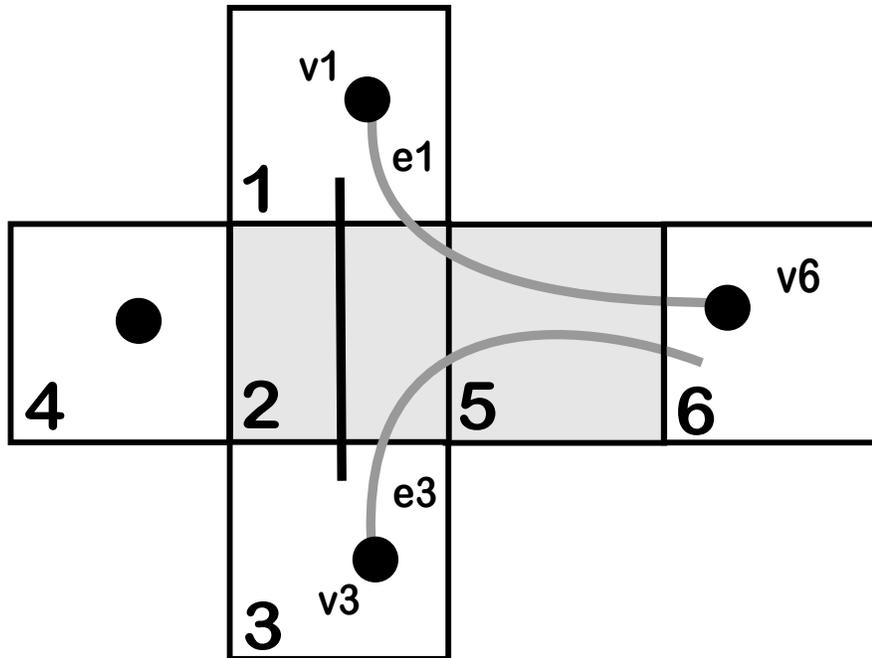


Figure 18.5: A vertical chain

In order to avoid finishing the proof as in Lemma 18.1 both edges e_1 and e_3 must cross $\Sigma_2 \cap \Sigma_5$. Neither edge can end in Σ_5 because Σ_5 is grid empty. If e_1 crosses the bottom edge of Σ_5 then e_1 blocks e_3 from exiting Σ_5 , which is a contradiction. Hence e_1 does not exit the bottom edge of Σ_5 . Similarly, e_3 cannot cross the top edge of Σ_5 . Hence, both e_1 and e_3 cross $\Sigma_5 \cap \Sigma_6$. But then e_1 and e_3 are errant edges, and this contradicts Statement 3 of the Pixellation Theorem.

18.4 Length Three Chains: Case C

Now we consider the remaining kind of length 3 chain, shown in Figure 18.6. As usual, let v_j be the graph grid point in Σ_j for $j = 1, 3$. We want to prove that an arithmetic graph edge connects v_1 to v_3 . This case is rather painful.

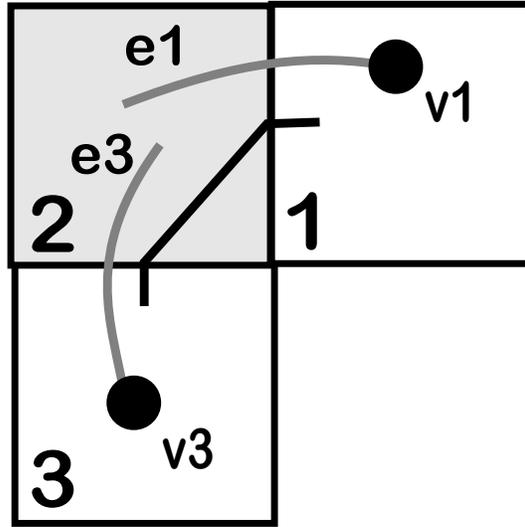


Figure 18.6

By Statement 5 of the Pixellation Theorem, $\Sigma_1 \cap \Sigma_2$ cannot be an unused edge with respect to Σ_1 , for the following reasons.

- If the associated offending edge were to cross through the top of Σ_1 , then the plaid segment in Σ_2 (which is part of the catch) would have the wrong position.
- If the associated offending edge were to cross through the bottom of Σ_1 , then the plaid segment in Σ_3 (which is part of the catch) would have the wrong position.

Hence some graph edge e_1 incident to v_1 crosses $\Sigma_1 \cap \Sigma_2$. The same argument applies to Σ_3 . So, we can assume that the edges e_1 and e_3 are as in Figure 18.6. We will show that e_1 has its other endpoint in Σ_3 . Since there is only one graph grid point in Σ_3 , the other endpoint of e_1 must be v_3 . In short, e_1 is the desired arithmetic graph edge connecting v_1 and v_3 . There are 6 situations we must rule out, and we deal with them in turn.

Case 1: Suppose e_1 ends in Σ_4 . Since $\Sigma_2 \cap \Sigma_4$ is not in the edge set of Σ_4 , the edge e_1 is offending with respect to Σ_4 . Given the negative slope of e_1 , the associated unused edge must be the right edge of Σ_4 and the catch must involve $\Sigma_1, \Sigma_2, \Sigma_4$. But then the plaid segment in Σ_1 is in the wrong position. This is a contradiction. The long Y-shaped graph in Figure 18.7 shows the shape of the plaid arc implied by the existence of the catch, and this contradicts the fact that the plaid arc in Σ_1 also crosses into Σ_2 .

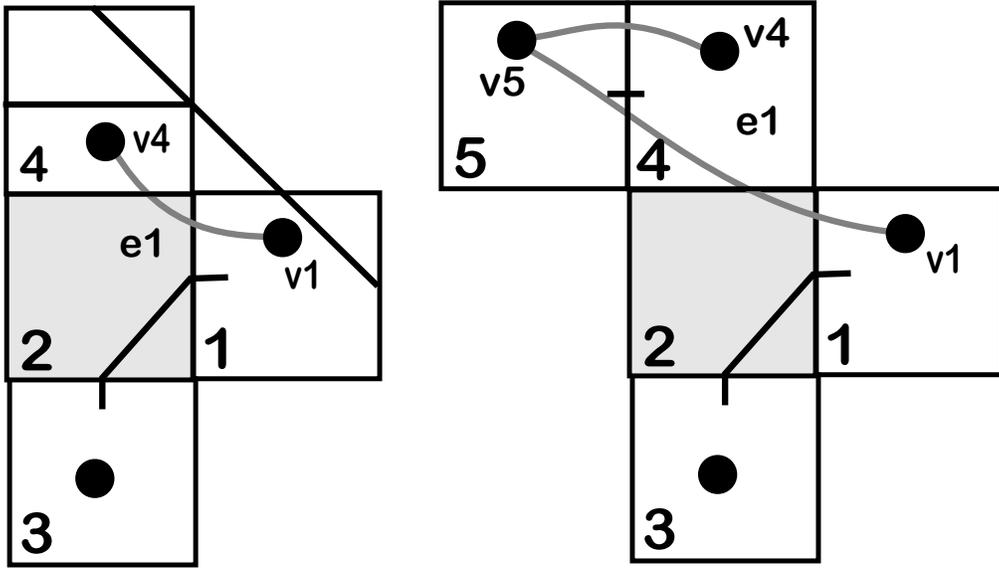


Figure 18.7: Cases 1 and 2

Case 2: Suppose e_1 ends in Σ_5 and that e_1 crosses $\Sigma_4 \cap \Sigma_5$. By Statement 4 of the Grid Geometry Lemma, the square Σ_4 is grid full. So, Σ_4 and Σ_5 are both grid full. If e_1 is an offending edge with respect to Σ_5 , then the associated unused edge is the bottom edge of Σ_5 . But then the catch for e_1 must involve $\Sigma_2, \Sigma_4, \Sigma_5$, and the plaid segment in Σ_2 is in the wrong position. Hence $\Sigma_4 \cap \Sigma_5$ is in the edge set for Σ_5 . Hence Σ_4, Σ_5 form a linked chain and some arithmetic graph edge joins v_4 and v_5 . This contradicts the Statement 4 of the Pixellation Theorem.

Case 3: Suppose that e_1 ends in Σ_5 and crosses $\Sigma_5 \cap \Sigma_6$, as shown on the left in Figure 18.8. If Σ_6 is not grid empty, then the same argument in Case 2 finishes the job. So, we may assume that Σ_6 is grid empty.

Now, e_1 cannot be an offending edge with respect to Σ_5 for the same reason as in Case 2. Thus, $\Sigma_5 \cap \Sigma_6$ is in the plaid edge set for Σ_6 . The left edge of Σ_6 cannot be in the plaid edge set for Σ_6 because this would make e_1 an errant edge for Σ_5 . Hence, the segment of the plaid model in Σ_6 connects $\Sigma_5 \cap \Sigma_6$ to $\Sigma_6 \cap \Sigma_7$. But then, by Case B in the previous section, some arithmetic graph edge connects Σ_5 to Σ_7 , as shown in Figure 18.10. But then two arithmetic graph edges cross $\Sigma_5 \cap \Sigma_6$, contradicting Statement 4 of the Pixellation Theorem.

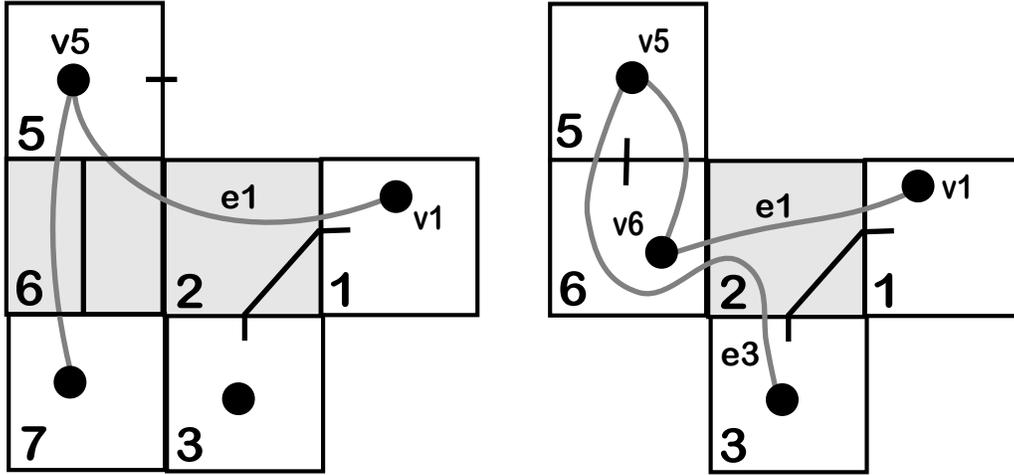


Figure 18.8: Cases 3 and 4

Case 4: Suppose that e_1 ends in Σ_6 and e_3 crosses $\Sigma_2 \cap \Sigma_6$. If e_3 ends at v_6 , then both e_1 and e_3 cross the same edge of Σ_6 and we contradict Statement 4 of the Pixellation Theorem. Since e_3 has length less than 2 and is not an errant edge for Σ_3 , the only possibility is that e_3 ends at v_5 . But the same argument as in Case 3 shows that e_3 is not an offending edge for Σ_5 . Hence Σ_6, Σ_5 is a linked chain. But then some other edge of the arithmetic graph connects v_6 to v_5 . This contradicts Statement 4 of the Pixellation Theorem.

Case 5: Suppose that e_1 ends in Σ_6 and e_3 does not cross $\Sigma_6 \cap \Sigma_2$. Note that e_1 blocks e_3 from crossing the top of Σ_2 , so e_2 crosses $\Sigma_1 \cap \Sigma_2$.

Note that e_1 is an offending edge for Σ_6 , because Σ_2 does not have $\Sigma_6 \cap \Sigma_1$ in its plaid edge set. The plaid segment in Σ_2 is not in the correct position for Σ_2 to be part of a catch for e_1 , as would happen if e_1 had negative slope. Hence e_1 has positive slope. Note that e_3 also has positive slope, because it rises up to cross $\Sigma_3 \cap \Sigma_2$. So, both e_1 and e_2 have the same slope. (In the reflected case, both would have negative slope.)

There are two kinds of distinguished edges having positive slope, and only one of them has a horizontal projection of length greater than 1. (The same goes in the negative slope case, which arises in the reflected case.) We know that the horizontal projection of e_1 is greater than 1. If the horizontal projection of e_3 is greater than 1, then e_1 and e_3 must be parallel. However, then we contradict Statement 6 of the Grid Geometry Lemma. We have found two parallel distinguished lines which intersect the same vertical edge of a square. In short e_3 has horizontal projection at most 1. Hence the endpoint of e_3 is either in Σ_1 or in Σ_7 .

If the endpoint of e_3 lies in Σ_1 , we are done. So, consider the case when the endpoint of e_3 lies in Σ_7 , as shown in Figure 18.9. In this case, we get the same contradiction as in Case 4.

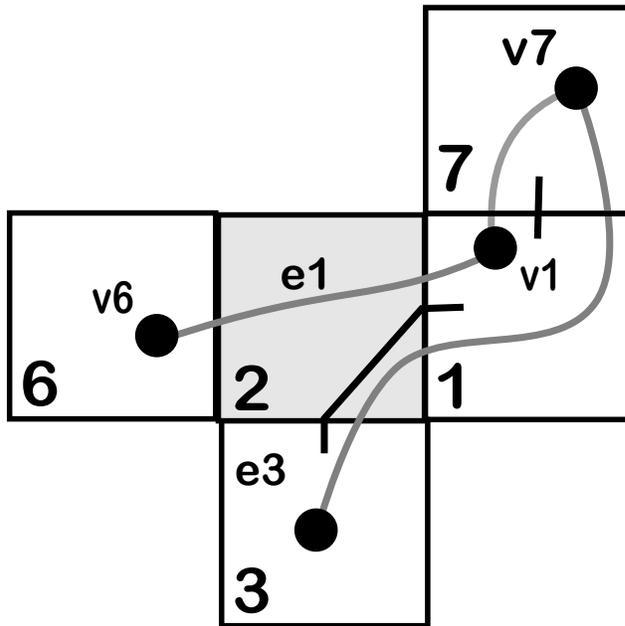


Figure 18.9: Case 5

Case 6: Suppose that $w_1 \in \Sigma_8$ or $w_1 \in \Sigma_9$. Then e_1 has positive slope and horizontal projection at least 1. If e_3 does not cross $\Sigma_2 \cap \Sigma_6$ then we get the same contradiction as in Case 5. So, we can assume that e_3 crosses $\Sigma_2 \cap \Sigma_6$, as shown in Figure 18.10. But then, as Figure 18.10 indicates, e_3 cannot be an offending edge for Σ_6 because the plaid segment in Σ_3 is not in the right position. Hence, $\Sigma_2 \cap \Sigma_6$ is in the edge set for Σ_6 . But this would make the edge set of Σ_2 have 3 members, which is a contradiction.

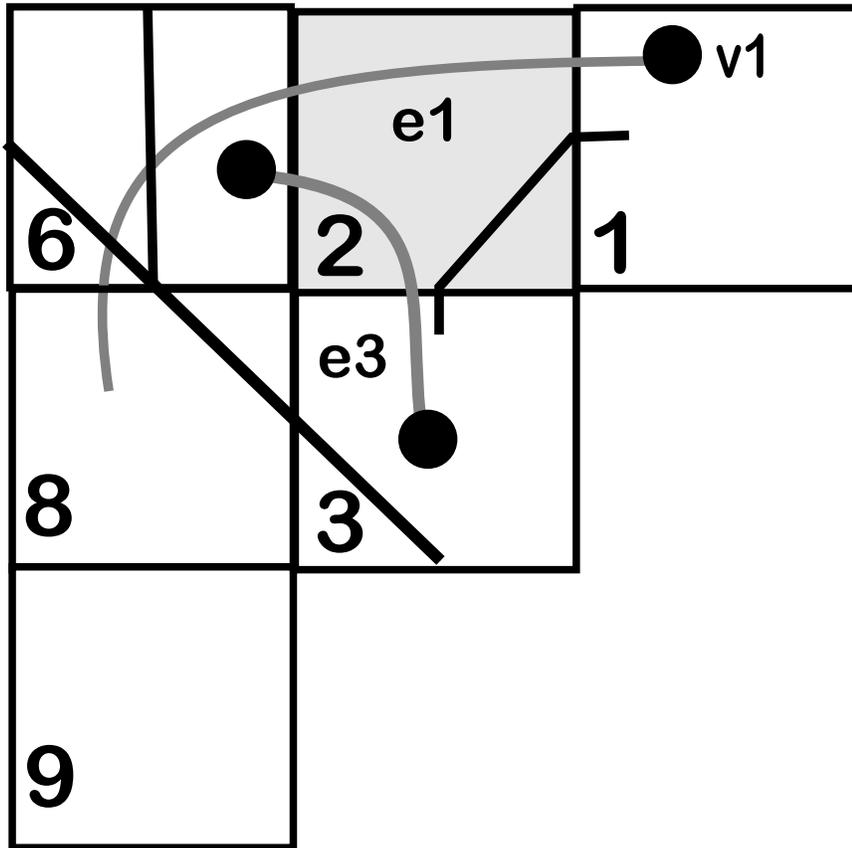


Figure 18.10: Case 6

This completes the analysis of Case C. At this point we have proved the Bound Chain Lemma for all linked chains, except those of length 4.

18.5 Length Four Chains: Case A

Suppose that $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$ is a length 4 linked chain, and all four of these squares are on the same row. The same argument as in Case 3A above shows that there must be some graph edge e_1 , incident to v_1 , which crosses $\Sigma_1 \cap \Sigma_2$. Figure 18.11 shows a tree of possibilities. We will show that, actually, this case cannot occur. Without loss of generality, we will consider the case when e_1 does not end up in a row of squares below our chain. Since e_1 has length less than 2, and e_1 is not an errant edge, e_1 must end either in Σ_5 or Σ_6 .

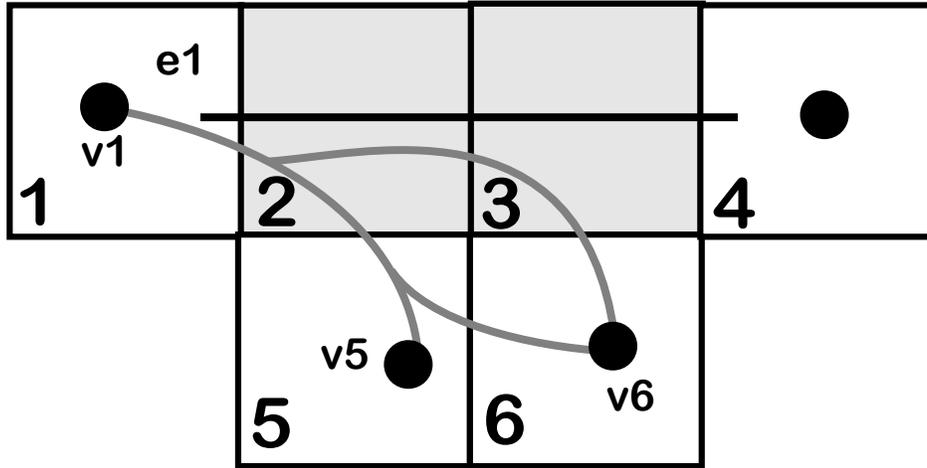


Figure 18.11: Four-in-a-row case

Case 1: Suppose e_1 lands in Σ_5 . Since $\Sigma_2 \cap \Sigma_5$ is an unused edge for Σ_5 and e_1 crosses this edge, e_1 must be an offending edge associated to the left edge of Σ_5 . But then $\Sigma_1, \Sigma_2, \Sigma_5$ are part of the catch for this edge. The plaid segment in Σ_1 is in the wrong position for this.

Case 2: Suppose e_1 lands in Σ_6 and crosses $\Sigma_5 \cap \Sigma_6$. Then e_1 cannot be an offending edge for Σ_6 because the plaid segment in Σ_3 is in the wrong position for it to participate in the required catch. Hence, some other arithmetic graph edge joins v_5 and v_6 . But then two graph edges are incident to v_6 and cross $\Sigma_5 \cap \Sigma_6$. This contradicts Statement 4 of the Pixellation Theorem.

Case 3: Suppose e_1 lands in Σ_6 and crosses $\Sigma_3 \cap \Sigma_6$. The same kind of argument as in Case 1 works here.

18.6 Length Four Chains: Case B

Suppose that $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$ is a length 4 linked chain, and exactly 3 of these squares are on the same row. Without loss of generality, we consider the case shown in Figure 18.12. The same argument as in Case A above shows that there must be some graph edge e_1 , incident to v_1 , which crosses $\Sigma_1 \cap \Sigma_2$.

Since e_1 has length less than 2 and cannot be an errant edge, e_1 must end in one of the squares $\Sigma_4, \Sigma_5, \Sigma_6, \Sigma_7$. Moreover, all these squares are grid full, by Statement 4 of the Grid Geometry Lemma. We want to show that e_1 ends in Σ_4 . There are 4 cases to rule out. All 4 cases are handled by arguments just like those in Cases 4A1 and 4A2 above.

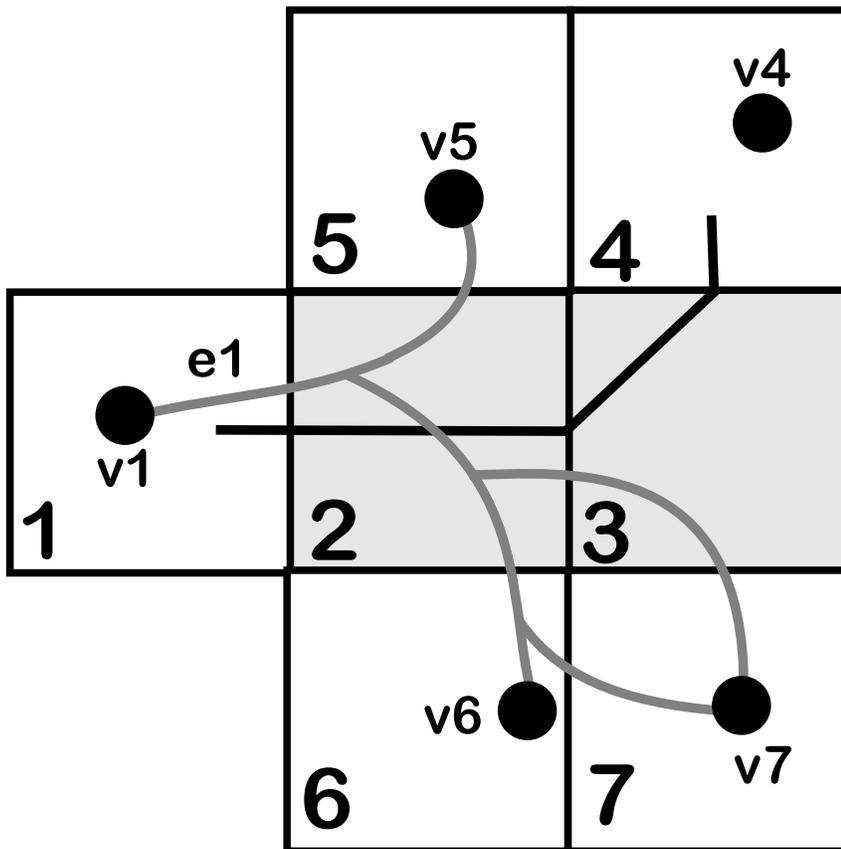


Figure 18.12: Three-in-a-row case.

Remark: We have stopped short of ruling out the existence of chains like this, but we think that they never actually occur.

18.7 Length Four Chains: Case C

Suppose that $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$ is a length 4 linked chain, making a 2×2 block as shown in Figure 18.13.

The same argument as in previous cases shows that there is some graph edge e_1 incident to v_1 which crosses $\Sigma_1 \cap \Sigma_2$. We want to show that e_1 connects v_1 to v_4 . Actually, this situation is impossible, given that we already know that e_1 crosses $\Sigma_1 \cap \Sigma_2$. So, our argument will really show that this kind of linked chain is impossible.

Before we start our analysis, we make special mention of the squares Σ_6 and Σ_9 . We do not know in general whether these squares are grid full or grid empty. However, these squares are irrelevant for all our arguments unless e_1 actually ends in them. In those cases, the relevant square is grid full, by definition. So, in all relevant cases, Σ_6 and Σ_9 are grid full, as drawn.

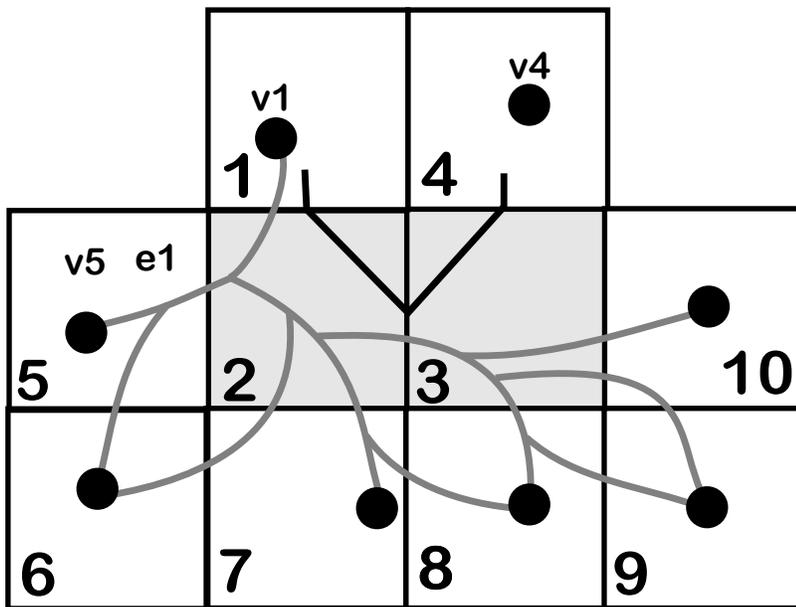


Figure 18.13: Block case

Case 1: Suppose e_1 ends in Σ_5 or Σ_{10} . This is the same as Case 4A1.

Case 2: Suppose e_1 lands in Σ_6 , Σ_8 , or Σ_9 . In all these cases, the argument is the same as in Case 4A2.

Case 3: Suppose that e_1 lands in Σ_7 . Here we redraw the picture to focus more particularly on this case. Since $\Sigma_7 \cap \Sigma_2$ is not in the edge set for Σ_2 , the edge e_1 must be an offending edge in Σ_7 associated to $\Sigma_7 \cap \Sigma_2$.

If e_1 has negative slope, then $\Sigma_2, \Sigma_5, \Sigma_6, \Sigma_7$ form the catch for e_1 . But the catch must be of the first kind, and e_1 must connect v_7 to v_5 . This is a contradiction. If e_1 has positive slope, then a similar argument shows that actually e_1 connects v_7 to v_4 , another contradiction.

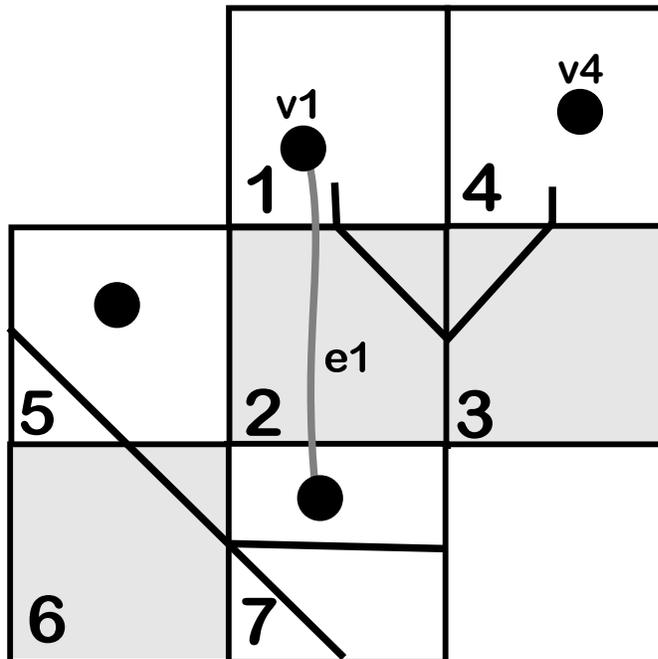


Figure 18.14: Case 3

18.8 Length Four Chains: Case D

There is only one remaining case, and this case actually occurs. Suppose that $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$ is a length 4 linked chain, making a zig-zag pattern as in Figure 18.15 below.

As in Case 3C, it could happen that the edge $\Sigma_1 \cap \Sigma_2$ is an unused side for Σ_1 . In this case, there is an offending edge e associated to $\Sigma_1 \cap \Sigma_2$. The catch for e must be of the second kind because Σ_3 is grid empty. The catch involves $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4, \Sigma_{10}$, and by definition e must end in Σ_4 as desired. So, we just have to worry about the case when some graph edge e_1 is incident to v_1 and crosses $\Sigma_1 \cap \Sigma_2$.

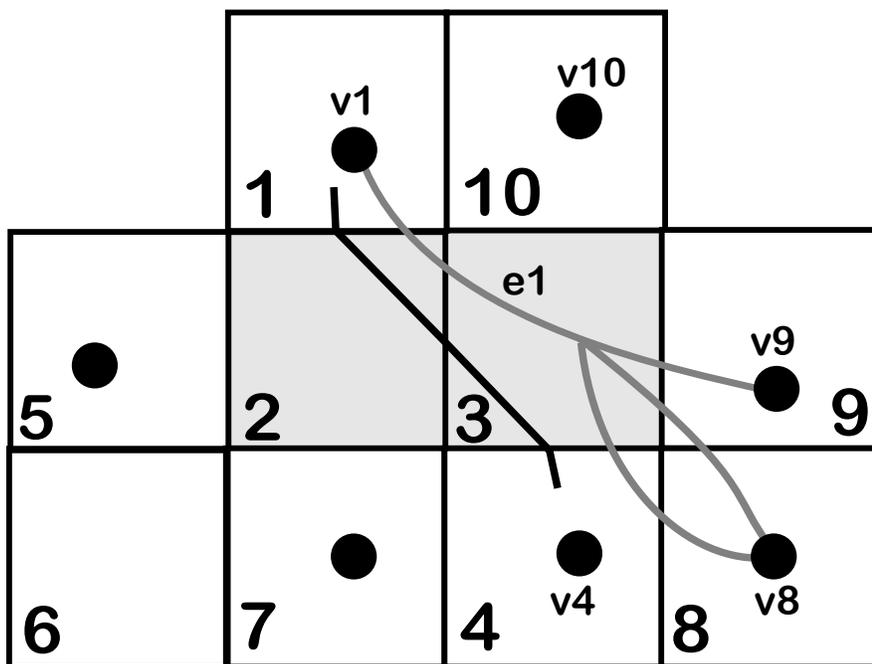


Figure 18.15: zig-zag case

The same reasons as in Case 4C rule out the possibility that e_1 ends in Σ_5 , Σ_6 , or Σ_7 . We just have to rule out e_1 ending in Σ_8 or Σ_9 . The same remarks as in Case 4C apply to the drawing of the square Σ_8 : This square could be either grid full or grid empty, but in the case relative to the argument it is grid full.

Case 1: Suppose e_1 ends in Σ_8 and crosses $\Sigma_4 \cap \Sigma_8$. The argument in Case 4B2 rules this out: Two graph edges would be incident to Σ_8 and would cross $\Sigma_4 \cap \Sigma_8$.

Case 2: Suppose e_1 ends in Σ_8 and crosses $\Sigma_9 \cap \Sigma_8$. If e_1 is not an offending edge for Σ_8 then we get the same contradiction as in Case 1. But if e_1 is an offending edge for Σ_8 then, inspecting the nature of the catches, we see that e_1 must connect v_8 to v_{10} .

Case 3: Suppose e_1 ends in Σ_9 . Looking at the portion of the plaid segment in Σ_3 , we see that $\Sigma_3 \cap \Sigma_9$ is not in the plaid edge set for Σ_9 . Hence e_1 is an offending edge for Σ_9 . The catch for e_1 must be of the first kind, because Σ_{10} is grid full. But then e_1 connects v_9 to v_{10} , a contradiction.

19 The Graph Reconstruction Formula

The purpose of this chapter is to present a formula which we call the *Reconstruction formula*, which relates the geometry of the graph grid to the graph PET classifying map. Combining this formula with the Graph Master Picture Theorem, we will prove Statement 1 of the Pixellation Theorem.

19.1 Main Result

Recall that $\widehat{X} = \mathbf{R}^3 \times [0, 1]$. For each parameter A , the affine group acting on \widehat{X} in connection with the graph PET acts as an abelian group of translations. Precisely, this group is $\Lambda(\mathbf{Z}^3)$, where

$$\Lambda = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 + A & 0 \\ -1 & 1 - A & 1 + A \end{bmatrix} \mathbf{Z}^3 \quad (174)$$

Given a point $\xi \in G$, let $[\xi] \in \mathbf{R}^2/\mathbf{Z}^2$ denote the equivalence class of ξ . We now explain how we can use the map Φ to determine $[\xi]$.

We introduce linear maps $\Theta_1, \Theta_2 : \mathbf{R}^3 \rightarrow \mathbf{R}$:

$$\Theta_1(x, y, z) = x, \quad \Theta_2(x, y, z) = \frac{y - Ax}{1 + A}. \quad (175)$$

Checking on the obvious basis for Λ , we see that $\Theta_j(\Lambda) \subset \mathbf{Z}$. Therefore Θ_1 and Θ_2 both give well defined maps from \mathbf{R}^3/Λ , the domain for the graph compactification at A , to the torus \mathbf{R}/\mathbf{Z} . Hence, we have a locally affine map

$$\Theta : \mathbf{R}^3/\Lambda \rightarrow \mathbf{R}^2/\mathbf{Z}^2, \quad L = [(\Theta_1, \Theta_2)]. \quad (176)$$

The goal of this section is to establish the following result.

Lemma 19.1 (Reconstruction)

$$[\xi] = \Theta \circ \Phi(\xi). \quad (177)$$

We call this equation the Reconstruction Formula. Equation 177 allows us to get control over how the arithmetic graph sits with respect to \mathbf{Z}^2 . We've already proven that no point of G lies on the boundary of an integer unit square. So, we can interpret both sides of Equation 177 as referring to points in the open integer unit square $(0, 1)^2$. We fix a parameter A throughout the proof.

Lemma 19.2 *Equation 177 holds at $\xi = T(0, 0)$.*

Proof: A direct calculation shows that

$$[\xi] = \left(\frac{1}{2q}, \frac{q-p}{2q(q+p)} \right).$$

Next, we compute

$$\Phi(\xi) = \left(\frac{1}{2q}, \frac{1}{2q}, \frac{1}{2q}, A \right).$$

A direct calculation then shows that

$$\Theta_2(\Phi(\xi)) = \frac{(p/q)(1/2q) - (1/2q)}{1 + (p/q)} = \frac{q-p}{2q(q+p)}.$$

Hence, the two sides of the equation agree. ♠

Lemma 19.3 *If Equation 177 holds at ξ , it also holds at $\xi \pm dT(0, 1)$.*

Proof: We do the (+) case. The (−) case has the same proof. Let $\xi' = \xi + dT(0, 1) = \xi + (1, -P)$. Here $P = 2A/(1 + A)$. We compute

$$\Theta \circ \Phi(\xi') - \Theta \circ \Psi(\xi) = \Theta(1, 1, 1) = \left(1, \frac{1-A}{1+A} \right) = (1, -P + 1).$$

The last expression is congruent to $\zeta' - \zeta = (1, -P) \pmod{\mathbf{Z}^2}$. ♠

Lemma 19.4 *If Equation 177 holds at ξ , it also holds at $\xi \pm dT(1, 1)$.*

Proof: We do the (+) case. The (−) case has the same proof. We have $dT(1, 1) = (1 + A, 1 + A)$. Using the same notation as in the previous argument, we have

$$\Theta \circ \Phi(\xi') - \Theta \circ \Psi(\xi) = \Theta(1 + A, 1 + A, 1 + A) = (1 + A, 1 - A) = \zeta' - \zeta.$$

This completes the proof ♠

These three lemmas together show that Equation 177 holds on all of $dT(\mathbf{Z}^2)$, as desired.

19.2 Eliminating Most Double Crossings

In this section we eliminate most of the double crossings mentioned in §17.3. This section does not use either the Reconstruction Formula or the Graph Master Picture Theorem.

Let T be the canonical affine transform. See §12.3. Recall that a *distinguished edge* is an edge having vertices $v_1, v_2 \in T(\mathbf{Z})$ such that $v_1 - v_2 = T(\zeta)$, where ζ is one of the 8 shortest nonzero vectors in \mathbf{Z}^2 . We say that the edge belongs to the family $\mathcal{F}(i, j)$ is $\zeta = \pm(i, j)$.

A *double crossing* is a configuration of the kind shown in Figure 19.1 below. The configuration consists of two adjacent unit integer squares Σ_1 and Σ_2 , graph grid points $v_1 \in \Sigma_1$ and $v_2 \in \Sigma_2$, and disjoint distinguished edges e_1 and e_2 incident to v_1 and v_2 respectively which both cross $\Sigma_1 \cap \Sigma_2$.

Lemma 19.5 *A bad configuration must have the following structure.*

- $\Sigma_1 \cap \Sigma_2$ is horizontal.
- The edge connecting v_1 to v_2 belongs $\mathcal{F}(1, 0)$.
- At least one of the two edges e_1 or e_2 belongs to $\mathcal{F}(-1, 1)$.

Proof: Let $f = v_2 - v_1$. By Statement 5 of the Grid Geometry Lemma, f must be a distinguished edge.

We define S_1 to be the set of 8 distinguished edges incident to v_1 . We think of these edges as vectors pointing out of v_1 . This set has a natural cyclic order on it. Likewise we define S_2 .

Let $\Sigma_{12} = \Sigma_1 \cap \Sigma_2$. The edges of S_1 which intersect Σ_{12} closest to $f \cap \Sigma_{12}$ are obtained by turning f one click in S_1 , either clockwise or counterclockwise. Likewise, the edges of S_2 which intersect Σ_{12} closest to $f \cap \Sigma_{12}$ are obtained by turning $(-f)$ one click in S_2 , counter clockwise or counterclockwise. The two turnings must be in the same direction, because otherwise the resulting edges would intersect.

So, the shortest possible distance between the two intersection points occurs when e_1 and e_2 are parallel. Statement 6 of the Grid Geometry Lemma now says that Σ_{12} is horizontal and the parallel lines are of type $(-1, 1)$. Looking at the proof of Statement 6 of the Grid Geometry Theorem given in §12.4, we see that this forced $f \in \mathcal{F}(1, 0)$. ♠

19.3 Eliminating the last Double Crossing

Now we will use the Reconstruction Formula and the Graph Master Picture Theorem to rule out the last kind of double crossing. This proves Statement 1 of the Pixellation Lemma. We will suppose that the arithmetic graph has a double crossing for some parameter A and derive a contradiction. The reader is warned in advance that our proof, at the end, is just a computer assisted calculation.

We will consider the case when the edge e_1 shown on the left side of Figure 19.1 is in the family $\mathcal{F}(-1, 1)$. The other case, when $e_2 \in \mathcal{F}(-1, 1)$, has a similar proof. Indeed, the second case follows from the first case and from the rotational symmetry of the arithmetic graph.

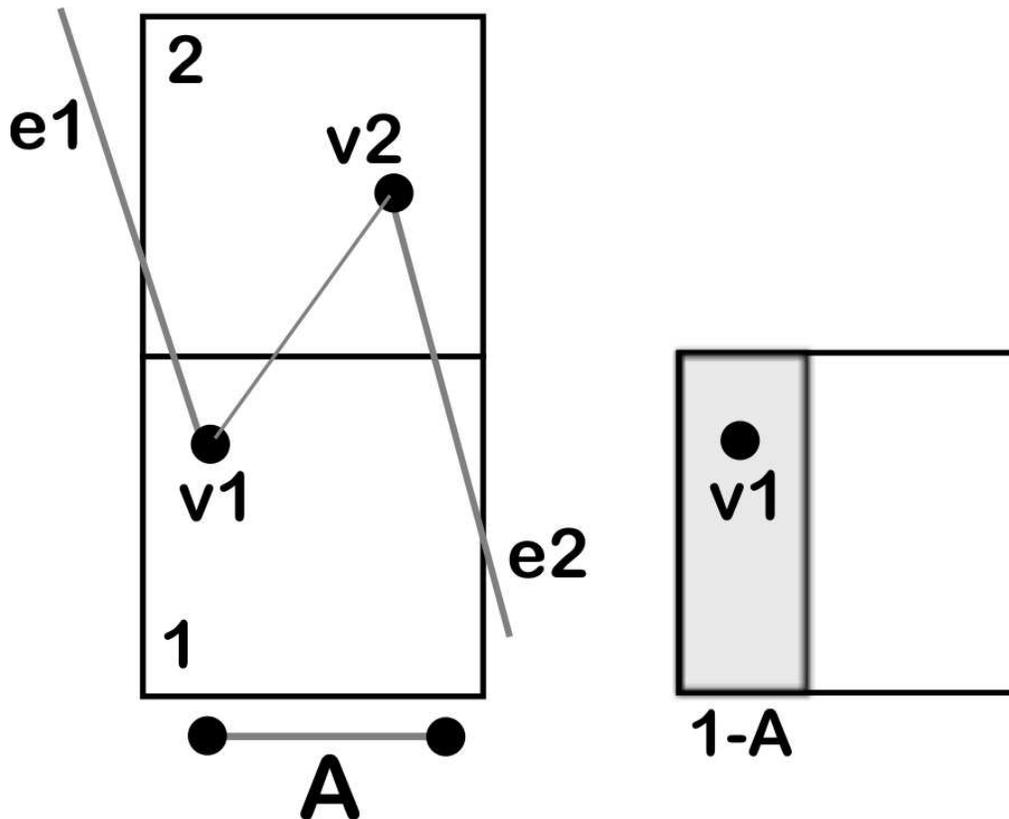


Figure 19.1: A double crossing.

Let G denote the graph grid and let $\Phi : G \rightarrow \widehat{X}$ be the graph classifying map.

Lemma 19.6 $\Phi(v_1)$ lies in the region of $\widehat{X} \cap H$, where H is the half space given by the equation $x < 1 - A$.

We know that the edge joining v_1 to v_2 is in the family $\mathcal{F}(1, 0)$. The first coordinate of $T(1, 0)$ is A , so the horizontal distance from v_1 to v_2 is A . That means that v_1 is within $1 - A$ units of the left edge of Σ_1 , as shown on the right half of Figure 19.1.

By the Reconstruction Formula, $\Phi(v_1)$ lies in the region

$$R_A = X_A \cap \{x \mid x < 1 - A\}. \quad (178)$$

R_A lies in the region

$$H \cap \widehat{X}, \quad (179)$$

where H is the halfspace given by the equation $x < 1 - A$. Recall that the coordinates on this space are (x, y, z, A) , so H is defined by integer equations. Also, H is invariant under the action of the graph lattice Λ defined in §13.1. ♠

Recall that we parametrize $\widehat{X} \subset \mathbf{R}^4$ with coordinates (x, y, z, A) . So, H is a halfspace defined by integer equations. Moreover, H is invariant under the action of the graph lattice defined in §13.1.

The way the graph classifying map works is that we look at $\Phi(v_1)$ and record the labels of the polytope in each partition which contains this point.

Lemma 19.7 *Assuming that a bad configuration exists, one of the polytopes in one of the partitions has the label $(1, -1)$ and intersects H .*

Proof: There are two polytopes of interest to us: the one in each partition which contains $\Phi(v_1)$. These polytopes both intersect H . One of these polytopes is responsible for the assignment of the vector e_1 to v_1 . Call this the *magic polytope*. Since e_1 is in the family $\mathcal{F}(-1, 1)$, the label of the magic polytope is either $(-1, 1)$ or $(1, -1)$. We check that $T(1, -1)$ is the one with positive y coordinate. Hence, the label of the magic polytope is $(1, -1)$. ♠

Now for the computer assisted part of the proof. We check that no polytope at all in the $(+)$ partition has the label $(1, -1)$, and the only polytopes in the $(-)$ partition intersect H . This is obvious from the pictures on my computer program, and I will describe in the last chapter the short linear algebra computation which checks it. This completes the proof.

20 The Hitset and the Intertwiner

20.1 The Hitset

We fix a parameter A . Recall that a unit integer square is *grid full* if it contains a graph grid point, and otherwise grid empty. Recall that Φ_{Π} is the plaid classifying map. Let G_{Π} denote the plaid grid - i.e., the centers of the unit integer squares. Let $G_{\Pi}^* \subset G_{\Pi}$ denote the set of centers of grid full squares. In this section we state a result which characterizes the image

$$\bigcup_{\zeta \in G_{\Pi}^*} \Phi_{\Pi}(\zeta). \quad (180)$$

The domain of Φ_{Π} is the unit cube

$$X_{\Pi} = [-1, 1]^3. \quad (181)$$

We define the *hitset* to be the subset of X_{Π} having the form

$$X_{\Pi}^* = H \times [-1, 1], \quad (182)$$

where H is the octagon with vertices

$$\begin{aligned} &(-1, 1), (-1 + P, -1 + P), (1 - P, -1), (1, -1 + P) \\ &(1, 1), (1 - P, 1 - P), (-1 + P, 1), (-1, 1 - P). \end{aligned} \quad (183)$$

The vertices are listed in cyclic order. The octagon H has a kind of zig-zag shape. Figure 9.1 shows the picture for several parameters.

In this chapter we will prove the following theorem.

Theorem 20.1 (Hitset) *For any parameter A , we have*

$$\bigcup_{\zeta \in G_{\Pi}^*} \Phi_{\Pi}(\zeta) \subset X_{\Pi}^*. \quad (184)$$

The Hitset Theorem is sharp in the following sense. All the objects make sense at irrational parameters as well as rational parameters, and for irrational parameters, the set on the left is dense in the set on the right. We will not prove this, because we do not need the result, but a proof would not be so difficult given everything else we prove in this chapter.

In Figure 20.1 below we show the polygon H for the parameters $k/5$ for $k = 1, 2, 3, 4$. (The picture makes sense for all parameters, and not just even rational ones.) Actually, we show not just H but also the image of H under the action of the lattice Λ_{Π} . We think of this lattice as acting on the xy plane just by forgetting about the third coordinate. The lattice is generated by the vectors $(2, P)$ and $(0, 2)$.

Thanks to the product structure of X_{Π} and X_{Π}^* , the planar pictures we show capture all the information. The various images of H fit together to form infinite bands which look like zigzags. As $A \rightarrow 0$ the zigzags come together and fill up the plane.

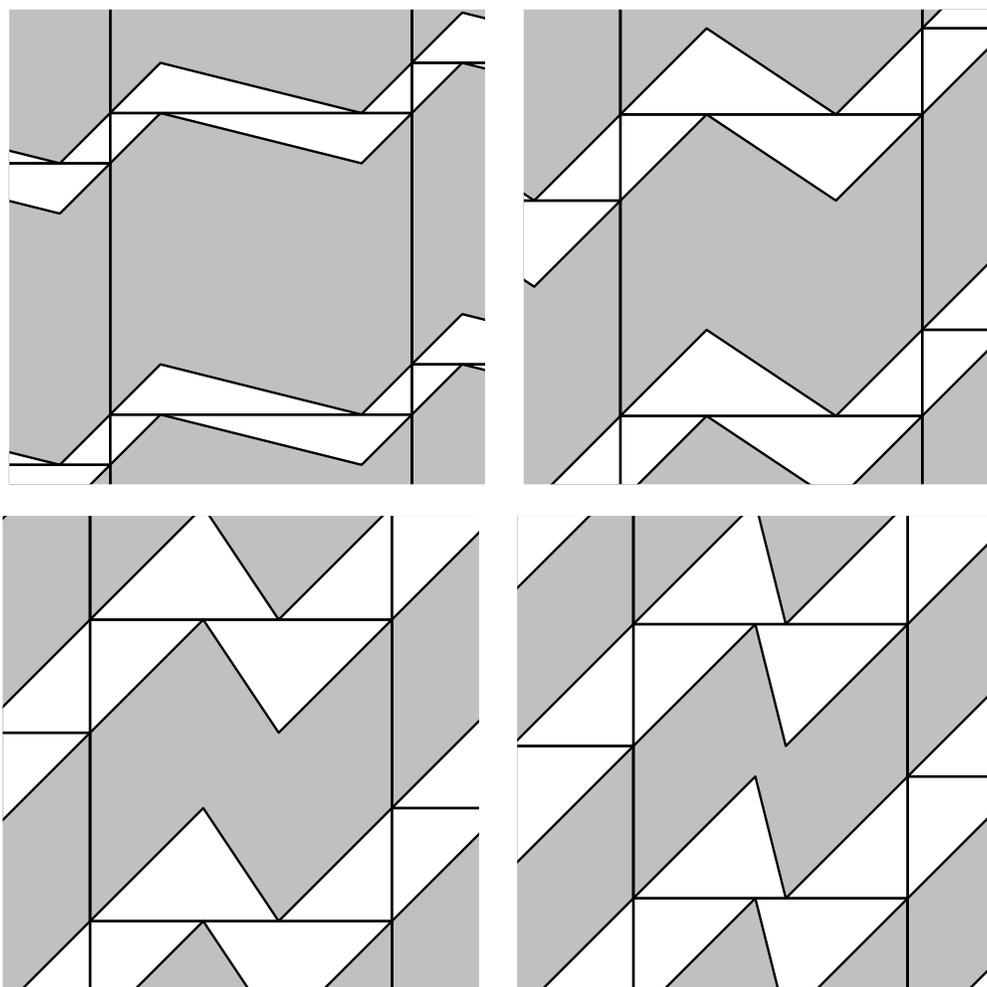


Figure 20.1: The orbit $\Lambda_{\Pi}(H)$ for parameters $A = k/5$, $k = 1, 2, 3, 4$.

20.2 The Projective Intertwiner

The symbol Ψ will denote the map which we call *the projective intertwiner*. We will define Ψ after we specify its domain and range. Recall that our total space is $\widehat{X} = \mathbf{R}^3 \times [0, 1]$.

Domain: The domain of Ψ is X_{Π} , though our theorem really only concerns the image of Ψ on the hitset X_{Π}^* . The set X_{Π} is the fundamental domain for the lattice Λ_1 acting on \widehat{X} . See §7.1.

We write

$$X_{\Pi} = X_{\Pi,-} \cup X_{\Pi,+}, \quad (185)$$

Where $X_{\Pi,+}$ consists of those points (x, y, z, P) where $x \leq y$ and $X_{\Pi,-}$ consists of those points there $y \leq x$. This is a partition of X_{Π} into two isometric halves.

Range: The range of Ψ is

$$X_{\Gamma} = \widehat{X}/\Lambda, \quad (186)$$

where Λ is the graph lattice defined in §13.1.

The Map: Now we define the map

$$\Psi : X_{\Pi} \rightarrow X_{\Gamma}.$$

For $(x, y, z, P) \in X_{\Pi,\pm}$, we define

$$\Psi(x, y, z, P) = \left[\frac{1}{2-P} \left(x - y, -y - 1, z + P + 1, P \right) \pm (1, 0, 0, 0) \right]_{\Lambda} \quad (187)$$

That is, we add or subtract 1 depending on which half of the partition our point lies in, and we take the result mod Λ . In the next section we check that Ψ is well defined even in boundary cases. The map Ψ is a piecewise defined integral projective transformation. If we hold P fixed and restrict Ψ to a slice, then Ψ is an affine transformation. We call Ψ the *projective intertwiner*.

The Main Result: Let G_{Γ} denote the graph grid. Let G_{Π} denote the plaid grid. Each point $\zeta_{\Gamma} \in G_{\Gamma}$ lies in the interior of a unique integer square. We let $\zeta_{\Pi} \in G_{\Pi}$ denote the center of this square.

Let $\Phi_{\Pi} : G_{\Pi} \rightarrow X_1$ denote the plaid classifying map. Recall that, at the parameter A , the range of Φ_{Π} is contained in the slice $\mathbf{R}^3 \times \{P\}$, where $P = 2A/(1 + A)$. Let $\Phi_{\Gamma} : G_{\Gamma} \rightarrow \widehat{X}/\Lambda$ be the graph classifying map.

Theorem 20.2 (Intertwining) *The following holds for every even rational parameter:*

$$\Phi_{\Gamma}(\zeta_{\Gamma}) = \Psi \circ \Phi_{\Pi}(\zeta_{\Pi}) \quad \forall \zeta_{\Gamma} \in G_{\Gamma}. \quad (188)$$

Remarks:

(i) I checked Equation 188 computationally for all relevant points and all parameter p/q with $q < 30$. This check is not meant as a substitute for a rigorous proof, but it is nice to know. I didn't check the Hitset Theorem as systematically, but my computer program plots the left and right hand sides of Equation 184, and one can see that it always works.

(ii) Notice in Equation 187 that the fourth coordinate on the right hand side is A , because $A = P/(2 - P)$. Thus, Ψ maps the relevant slices to each other.

(iii) It might be nicer if there were a global projective transformation from \widehat{X} to itself which works in place of our piecewise projective map Ψ . However, we have $\Psi(T_X(V) - V) = (2, 0, 0, 0)$. The vector on the right does not have for them $\lambda(W) - W$ for any transformation $\lambda \in \Lambda_1$, the plaid lattice. This situation makes the existence of a global projective intertwiner impossible. On the other hand, below in §20.8 we will modify the domain of Ψ so that Ψ is projective throughout the interior of the new domain. When we look at the action of Ψ on the new domain, we will see the canonical nature of Ψ . See Lemma 20.4 for instance.

(iv) When it comes time to prove the Pixellation Theorem, we shall be interested in the action of Ψ on the polytopes of the plaid triple partition. At first, it looks like we might have trouble, due to the piecewise nature of Ψ . However, it turns out that every plaid triple polytope is contained in one of the two pieces of the partition of X_{Π} .

(v) Our proof will show that a suitable formulation of the Intertwining Theorem holds for all parameters, and not just even rational ones. Indeed, it basically follows from continuity.

(vi) The reader who is keen to see how the proof of the Pixellation Theorem works might want to take the Hitset Theorem and the Intertwining Theorem for granted on the first reading. Our proof of the Pixellation Theorem only uses the truth of these statements, and not any theory developed during their proof. §21.3 gives a good illustration of how we use the Intertwining Theorem.

20.3 Well Definedness

Since Ψ is only piecewise defined, we have to worry about the cases when there are two competing definitions for Ψ . In our discussion of this matter, the symbol $(*)$ stands for a coordinate value that we don't care about. We fix some even rational parameter for the discussion.

There are three issues. One issue is that perhaps

$$\Phi_{\Pi}(\zeta) = (-1, *, *).$$

In this case, the two points

$$\phi_+ = \Phi_{\Pi}(\zeta), \quad \phi_- = \Phi_{\Pi}(\zeta) + (2, P, P)$$

are equally good representatives. Here $\phi_{\pm} \in X_{\Pi, \pm}$. An easy calculation shows that

$$\begin{aligned} \Psi(\phi_-) - \Psi(\phi_+) &= \frac{1}{2-P} (2-P, -P, PP, -P, 2-P) - (2, 0, 0) = \\ &(-1, -A, A) \in \Lambda_1. \end{aligned} \tag{189}$$

Hence, either representative gives rise to the same point in the range.

The second issue seems more serious, in view of Remark (v) above. It might happen that

$$\phi = \Phi_{\Pi}(\zeta) = (v, v, *).$$

That is, ϕ lies on the boundary of both pieces of the partition of X_{Π} . Let's check that this situation cannot arise for the images of points of G_{Π} . Such points have the form (x, y) where x and y are half integers. We compute that

$$\Phi_{\Pi}(x, y) = (2Px + 2y, 2Px, *) + (2m, Pm, *), \tag{190}$$

for a suitable integer m . We also observe that $2y$ is odd. Equation 190 leads to

$$P = \frac{2m + 2y}{m}.$$

This is impossible, because $P = 2p/(p+q)$, and $2m + 2y$ is an odd integer.

The third issue is that $\Phi_{\Pi}(\zeta) = (*, u, v)$ with either $u = \pm 1$ or $v = \pm 1$ or both. The case $u = \pm 1$ is impossible for similar reasons that we have just discussed. When $v = \pm 1$ it has replacing v by $v \mp 2$ has no effect on the Intertwining formula.

20.4 Strategy of the Proof

The rest of the chapter is devoted to proving the Hitset Theorem and the Intertwining Theorem. We prove these two results at the same time because they are closely related to each other. Here are the steps of the proof. We fix a parameter $A = p/q$.

- We prove the Intertwining Theorem for points of the form

$$\zeta_n = (n + 1/2)(1 + A, 1 - A), \quad n = 0, 1, 2, \dots \quad (191)$$

These points all belong to G_Γ because ζ_0 is the anchor point and $\zeta_n - \zeta_0 = ndT(1, 1)$. Here T is the canonical affine transformation. We call the points in Equation 191 the *diagonal points*.

- We prove the Hitset Theorem for the points in Equation 191.
- We prove the following induction step: Suppose that the Hitset Theorem is true for some graph grid point ζ . Then it is also true for $\zeta + dT(0, 1)$. We call this *Hitset Induction*.
- We prove the following induction step: Suppose that the Intertwining Theorem is true for some graph grid point ζ . Then it is also true for $\zeta + dT(0, 1)$. We call this *Intertwiner Induction*.

Let L denote the lattice of symmetries generated by the vectors $(\omega^2, 0)$ and $(0, \omega)$. Here $\omega = p + q$, as usual. We have already shown that both the plaid model and the arithmetic graph are invariant under L . We say that an *L-orbit* is an orbit of L acting on the graph grid. By symmetry, it suffices to prove our two theorems on a set which contains at least one point of every L -orbit

The four steps above combine to to prove that the two theorems hold true on sets of the form $dT(B)$ where B is a ball in \mathbf{Z}^n of arbitrarily large radius. For sufficiently large B , the set $dT(B)$ intersects every L orbit.

Remark: It would have been nice if the set of points $\{\zeta_n\}$ intersected every L -orbit. If this was true, we would not need the induction part of the proof. Likewise, it would have been nice if, starting from a single point (such as the anchor), we could reach every L -orbit just by repeatedly adding $dT(0, 1)$. If this was true, we would not need the first two steps of our proof. Alas, neither of these step-saving situations is true.

20.5 The Intertwining Theorem on the Diagonal

In this section we prove the Intertwining Theorem for points of the form

$$\zeta_n = \left(n + \frac{1}{2}\right)(1 + A, 1 - A), \quad n = 0, 1, 2, \dots \quad (192)$$

These points all belong to G_Γ , because $\zeta_n = \zeta_0 + dT(n, n)$ and ζ_0 is the anchor point.

We fix the value of n . It is convenient to consider the sub-intervals

$$R_{n,k} = \left(\frac{2k}{2n+1}, \frac{2k+1}{2n+1}\right), \quad k = 0, \dots, (n-1). \quad (193)$$

$$L_{n,k} = \left(\frac{2k-1}{2n+1}, \frac{2k}{2n+1}\right), \quad k = 1, \dots, (n-1). \quad (194)$$

We ignore the endpoints of these intervals; the boundary cases, when relevant, follow from continuity.

Let $I = (A_0, A_1)$ be one of the intervals of interest. We will sometimes use the following trick to bound certain numerical quantities that depend on $A \in I$. When the quantity is monotone, we get the bounds by evaluating the expression on the boundary values $A = A_0$ and $A = A_1$. We call this method *the boundary trick*.

When $A \in L_{n,k}$ (respectively $R \in L_{n,k}$, the point ζ_n lies in the left (respectively right) half of the square with center

$$\zeta_{n,k} = \left(n + \frac{1}{2}, n + \frac{1}{2}\right) + (k, -k). \quad (195)$$

First we consider the case when $A \in L_{n,k}$. We have

$$\Phi_\Pi(\zeta_{n,k}) \equiv (P(2n+2k+1) + (2n-2k+1), P(2n+2k+1), 2P(2n+1)). \quad (196)$$

The symbol \equiv means that we still need to reduce mod Λ_1 to get a vector in the fundamental domain.

The boundary trick tells us that the first coordinate in Equation 196 lies in

$$\left(1 + 2k + 2n - \frac{1 + 2n}{k + n}, 1 + 2k + 2n\right) \quad (197)$$

So, we subtract off the lattice vector $(n+k)(2, P, P)$. This gives

$$\Phi_\Pi(\zeta_{n,k}) \equiv (P(2n+2k+1) - 4k + 1, P(1+k+n), 2P(2-k+3n)). \quad (198)$$

The boundary trick tells us that the second coordinate in Equation 198 lies in the interval $(-1, 1) + 2k$. Subtracting $(0, 2k, 0)$ from Equation 198, we get a point in the fundamental domain:

$$\Phi_{\Pi}(\zeta_{n,k}) = (P(2n+2k+1)-4k+1, P(1+k+n)-2k, P(2-k+3n)+2\beta). \quad (199)$$

Here β is some integer whose value we don't care about.

The interval check shows that the first coordinate in Equation 199 is larger than the second coordinate. (We apply the trick to the difference of the coordinates.) Hence the point in Equation 199 lies in $X_{\Pi,-}$. We also have

$$\Phi_{\Gamma}(\zeta) \equiv \left(n + \frac{1}{2}\right)(1 + A, 1 + A, 1 + A) \bmod \Lambda. \quad (200)$$

Here Λ is the graph lattice. (In all these equations we are leaving off the fourth coordinate; we know this works out already.)

A calculation, which we do in Mathematica, shows that

$$\begin{aligned} \Psi(\Phi_{\Pi}(\zeta_{n,k}) - \Phi_{\Gamma}(\zeta) = & \\ & (-1 - k - n)(1, -1, -1) + \\ & (-2 - 2n)(0, 1 + A, 1 - A) + \\ & (1 - k + \beta)(0, 0, 1 + A). \end{aligned}$$

In other words, we have written the difference between the quantities as an integer combination of vectors in the graph lattice Λ . Hence, the two quantities are equal, as desired.

When $A \in R_{n,k}$ the calculation is very similar and we will just describe the differences. This time the interval trick tells us to subtract off the vector $(n + k + 1)(2, P, P)$ from the point in Equation 196. Then, as in the other case, we subtract off $(0, -2k, 0)$. The result is

$$\Phi_{\Pi}(\zeta_{n,k}) = (P(2k+2n+1)-4k-1, P(k+n)-2k, P(1-k+3n)+\beta) \quad (201)$$

The interval trick shows that this point lies in $X_{\Pi,+}$. Now we compute

$$\begin{aligned} \Psi(\Phi_{\Pi}(\zeta_{n,k}) - \Phi_{\Gamma}(\zeta) = & \\ & (-k - n)(1, -1, -1) + \\ & (-1 - 2n)(0, 1 + A, 1 - A) + \\ & (1 - k + \beta)(0, 0, 1 + A). \end{aligned}$$

This gives us the same desired conclusion as in the first case.

20.6 The Hitset Theorem on the Diagonal

Now we prove the Hitset Theorem for the points we considered in Step 1. We first consider the case when $A \in L_{n,k}$. The triangle Δ_L with vertices

$$(-1 + P, -1 + P), \quad (1, -1 + P), \quad (1, 1) \quad (202)$$

is the convex hull of 3 of the vertices and is contained in the polygon defining the hitset.

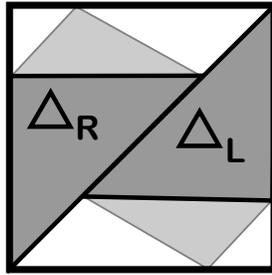


Figure 20.2: The triangles Δ_L and Δ_R .

Let ξ denote the point in the plane obtained by taking the first two coordinates of the point in Equation 199. That is:

$$\xi = (P(2n + 2k + 1) - 4k + 1, P(1 + k + n) - 2k). \quad (203)$$

It suffices to prove that $\xi \in \Delta_L$ for all $A \in L_{n,k}$. One of the sides of Δ_L is the line $x = 1$, and certainly ξ does not cross this line; it corresponds to one of the sides of the fundamental domain. Another side of Δ_L is the line $y = -1 + P$. The interval trick shows that ξ stays above this line. Finally, the other side of Δ_L is the line $y = x$. We already know that $\Phi_{\Pi}(\zeta) \in X_{\Pi,-}$, and this is the equivalent to the statement that ξ lies to the right of this line. These three conditions together imply that $\xi \in \Delta_L$. My computer program allows you to see a plot of ξ and Δ_L for all the relevant parameters.

Now consider the case when $A \in R_{n,k}$. This time we use the triangle $\Delta_R = -\Delta_L$, obtained by negating all the coordinates of the vertices of Δ . This time we want to show that the point

$$(P(2k + 2n + 1) - 4k - 1, P(k + n) - 2k) \quad (204)$$

lies in Δ_R . The same arguments as above shows that this point lies to the right of the line $x = -1$, below the line $y = 1 - P$, and to the left of the line $y = x$. This does the job for us.

20.7 Hitset Induction

Here we will prove Hitset Induction modulo what we call *the geometric claim*. We will prove the geometric claim at the end of the chapter.

We set $\zeta_\Gamma = \zeta$ and $\zeta'_\Gamma = \zeta + dT(0, 1)$. Similarly, we define ζ_Π and ζ'_Π . Here ζ_Π is the center of the unit integer square that contains ζ_Γ . Since we will be mentioning both the plaid lattice and the graph lattice, we let Λ_Π denote the plaid lattice and Λ_Γ denote the graph lattice.

We have

$$dT(0, 1) = (1, -P). \quad (205)$$

From this equation, we see that we have one of two possibilities for $\zeta_{\Gamma'} - \zeta_\Pi$. This difference either equals $(1, 0)$ or $(1, -1)$. To be more precise, let G_Γ^{hi} denote the union of those choices of $\zeta_\Gamma = (x, y)$ such that $x - \text{floor}(x) > P$. Define ζ_Π^{lo} to be the complementary set. For even rational parameters, it never happens that $y = \text{floor}(y) = P$, because then ζ'_Γ would be lie on the boundary of a unit integer square.

Let H^{lo} denote the parallelogram with vertices

$$(1-3P, 1-3P), \quad (1-P, 1-P), \quad (-1+P, 1) \quad (-1-P, 1-2P). \quad (206)$$

Let H^{hi} denote the polygon with vertices

$$(-1+P, -1+P), \quad (1-P, -1), \quad (3-3P, 1-2P) \quad (1-P, 1-P). \quad (207)$$

These polygons are not subsets of H . However, the orbits $\Lambda_\Pi(H^{\text{hi}})$ and $\Lambda_\Pi(H^{\text{lo}})$ together give a partition of $\Lambda_\Pi(H)$. The dark region is the *lo orbit* and the light region is the *hi orbit*.

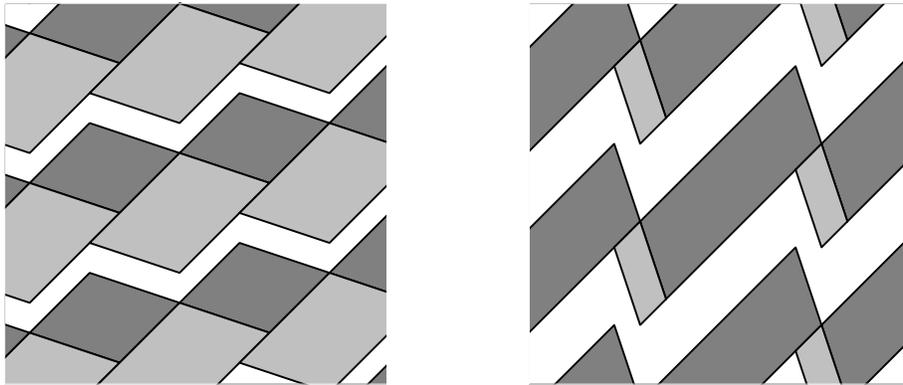


Figure 20.4: The orbits $\Lambda_\Pi H^{\text{lo}}$ and $\Lambda_\Pi H^{\text{hi}}$ for $A = 1/4$ and $A = 3/4$.

Geometric Claim:

- $\zeta_\Gamma \in G_\Gamma^{\text{hi}}$ implies $\Phi_\Pi(\zeta_\Pi) \in \Lambda_\Pi(H^{\text{hi}} \times [-1, 1])$
- $\zeta_\Gamma \in G_\Gamma^{\text{lo}}$ implies $\Phi_\Pi(\zeta_\Pi) \in \Lambda_\Pi(H^{\text{lo}} \times [-1, 1])$

We call this claim the *geometric claim*.

We introduce the new sets

$$(H^{\text{hi}})' = H^{\text{hi}} + (2P, 2P), \quad (H^{\text{lo}})' = H^{\text{lo}} + (2P - 2, 2P - 2). \quad (208)$$

If $\zeta_\Gamma \in G_\Gamma^{\text{hi}}$ and the geometric claim is true, then

$$\Phi_\Pi(\zeta'_\Pi) \in \Lambda_\Pi((H^{\text{hi}})' \times [-1, 1]).$$

The same goes when we replace *hi* with *lo*. In the low case, we are using the fact that $(2P - 2, 2P - 2)$ and $(2P - 2, 2P)$ are equal up to a vector in Λ_Π .

Now for the punchline. $\Lambda_\Pi((H^{\text{lo}})')$ and $\Lambda_\Pi((H^{\text{hi}})')$ give a second partition of $\Lambda_\Pi(H)$. Figure 20.5 shows the picture for the parameters $A = 1/4$ and $A = 3/4$.

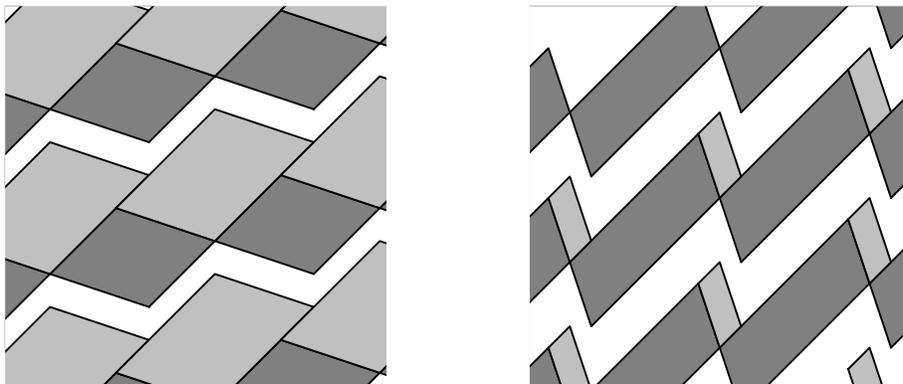


Figure 20.5: The orbits $\Lambda_\Pi(H^{\text{lo}})'$ and $\Lambda_\Pi(H^{\text{hi}})'$ for $A = 1/4$ and $A = 3/4$.

We get exactly the same picture as in Figure 20.4 except that the pieces have each been translated.

What we are really saying is that there is an infinite polygon exchange transformation on $\Lambda_\Pi(H)$ which corresponds to the operation of adding the vector $dT(0, 1)$ in G_Γ . The orbit $\Lambda_\Pi(H)$ decomposes into a countable union of parallelograms, each partitioned into a light parallelogram and a dark one. Our map simply exchanges the light and dark pieces within each component. On each component, our map is essentially a rotation of a flat torus.

Figure 20.6 shows the two partitions side by side, for the parameter $A = 1/4$. The “components” we are talking about are parallelograms which are bounded on opposite sides by lines of slope 1. For later reference, we call these parallelograms *dipoles*. So, again, each dipole is partitioned into a light and dark parallelogram, and our polygon exchange simply exchanges the two pieces within each dipole.

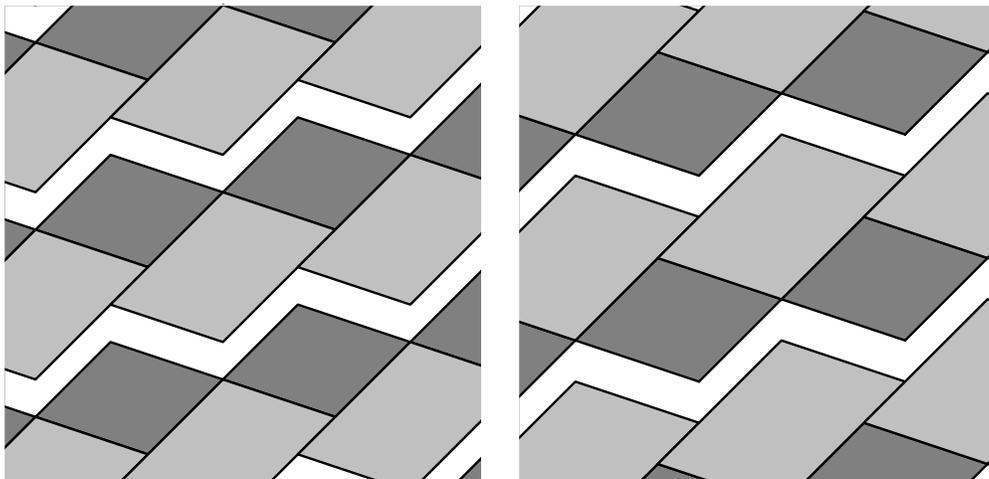


Figure 20.6: The two partitions of $\Lambda_{\Pi}(H)$ for the parameter $A = 1/4$.

The Geometric Claim immediately implies Hitset Induction. Hence, the Geometric Claim implies the Hitset Theorem.

20.8 Changing the Fundamental Domain

Observe that the orbit $\Lambda_{\Pi}(H)$ is simply a union of dipoles. Rather than consider $X_{\Pi} = [-1, 1]^3$ as our fundamental domain for the action of Λ_{Π} , we instead consider the fundamental domain to be

$$\Upsilon \times [-1, 1] \tag{209}$$

where Υ is the dipole that intersects $[-1, 1]^2$ to the right of the diagonal line $y = x$. Figure 20.7 shows Υ and how it sits with respect to $[-1, 1]$.

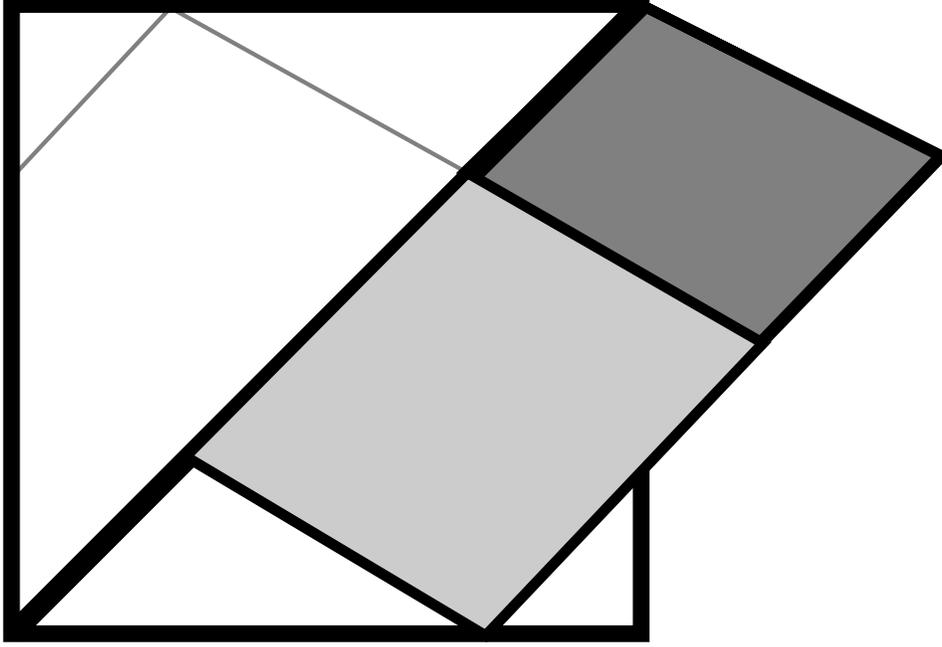


Figure 20.7: The fundamental dipole.

We get an easier calculation if we use $\Upsilon \times [-1, 1]$ as the domain for the projective intertwiner Ψ . To do this, we need to define Ψ on the whole fundamental domain, and check that the new definition agrees with the original in the appropriate sense.

The points of $\Upsilon - [-1, 1]^2$ all lie to the right of $[-1, 1]^2$, and one simply subtracts off $(-2, -P)$ to get them back into $[-1, 1]^2$. For any

$$\phi \in \left(\Upsilon - [-1, 1]^2 \right) \times [-1, 1], \quad (210)$$

We define $\Psi(\phi)$ using the branch of Φ which is defined for $X_{\Pi,-}$. Let $\phi^* = \phi - (2, P, P) \in X_{\Pi,+}$. Using the fact that $\phi \in X_{\Pi,-}$ and $\phi^* \in X_{\Pi,+}$, we compute

$$\Psi(\phi) - \Psi(\phi^*) = (-1, -A, A) = (-1, 1, 1) - (0, 1 + A, 1 - A) \in \Lambda_{\Gamma}. \quad (211)$$

So, our redefinition does not change anything. We can prove Induction Statement 2 using the new fundamental domain. Moreover, the restriction of Ψ to the new domain is projective throughout.

20.9 Intertwiner Induction

Since $\Phi_\Gamma(x, y) = (x, x, x) \bmod \Lambda_\Gamma$, we have

$$\Phi_\Gamma(\zeta'_\Gamma) - \Phi_\Gamma(\zeta_\Gamma) = (1, 1, 1) \bmod \Lambda_\Gamma. \quad (212)$$

Let

$$\phi = \Phi_\Pi(\zeta_\Pi), \quad \phi' = \Phi_\Pi(\zeta'_\Pi). \quad (213)$$

To establish Intertwiner Induction, we just have to prove

$$\Psi(\phi') - \Psi(\phi) - (1, 1, 1) \in \Lambda_\Gamma. \quad (214)$$

Using our new fundamental domain, and also the action of the map Φ_Π , we have one of

$$\phi - \phi' = (2P, 2P, 2P + 2\beta), \quad \phi' - \phi = (2P - 2, 2P - 2, 2\beta), \quad (215)$$

depending on whether we are in the hi case or the lo case. Here β is some integer whose value does not effect the calculation.

Hi Case:

$$\begin{aligned} \Psi(\phi') - \Psi(\phi) - (1, 1, 1) &= \\ \frac{1}{2 - P} \left(0, -2P, 2P + 2\beta \right) - (1, 1, 1) &= \\ (-1, -2A - 1, 2A - 1 + 2\beta(1 + A)) &= \\ 2\beta(0, 0, 1 + A) + (-1, 1, 1) + 2(0, A + 1, A - 1) &\in \Lambda_\Gamma. \end{aligned}$$

Notice that β plays no role at all in the final answer. To simplify the second calculation, we assume that $\beta = 0$.

Lo Case:

$$\begin{aligned} \Psi(\phi') - \Psi(\phi) - (1, 1, 1) &= \\ \frac{1}{2 - P} \left(0, -2P + 2, 0 \right) - (1, 1, 1) &= \\ (-1, -A, -1) = (-1, 1, 1) - (0, 1 + A, 1 - A) - (0, 0, 1 + A) &\in \Lambda_\Gamma. \end{aligned}$$

This is what we wanted to prove.

At this point, we have reduced the Intertwining Theorem and the Hitset Theorem to the Geometric Claim.

20.10 Proof of the Geometric Claim

We will combine the Intertwining Formula with the Graph Reconstruction Formula. We are allowed to do this for the pair $(\zeta_\Gamma, \zeta_\Pi)$ by induction.

We introduce coordinates

$$\zeta_\Gamma = (a, b), \quad \Phi_\Pi(\zeta_\Pi) = (x, y, z, P)$$

. Let $[t] = t - \text{floor}(t)$.

Lemma 20.3

$$[a] = \left[\frac{x - y}{2 - P} \right], \quad [b] = \left[\frac{-2 - P + P^2 + Px + 2y - 2Py}{2P - 4} \right]. \quad (216)$$

Proof: Define

$$\Psi \circ \Phi_\Pi(\zeta_\Pi) = (x^*, y^*, z^*, A).$$

For convenience, we repeat Equation 187 here:

$$\Psi(x, y, z, P) = \left[\frac{1}{2 - P} (x - y, -y - 1, z + P + 1, P) - (1, 0, 0, 0) \right]_\Lambda \quad (217)$$

We always take the $(-)$ option in Equation 187 because we are using the domain $\Upsilon \times [-1, 1]$ described in §20.8. The graph reconstruction formula tells us that

$$[a] = [x^*], \quad [b] = \left[\frac{y^* - Ax^*}{1 + A} \right].$$

Combining this formula with Equation 217 and doing some algebra, we get Equation 216. ♠

The next result says that the formulas in Equation 216, which look messy, are actually as nice as possible.

Lemma 20.4 *Equation 216 induces an affine diffeomorphism from Υ to $[0, 1]^2$.*

Proof: Forgetting about the brackets, the corresponding map on the plane is an affine diffeomorphism Ω . The first coordinate of Ω is obviously constant along lines of slope 1. These lines are parallel to the diagonal sides of Υ . We

claim that the second coordinate of Ω is constant along lines parallel to the other two sides of Υ . To see this, we plug in the equation

$$y = \left(\frac{-P}{2-P} \right) x + c. \quad (218)$$

and observe that the resulting expression

$$f(c) = \frac{2+P-P^2}{4-2P} + \left(\frac{P-1}{2-P} \right) c \quad (219)$$

is independent of x .

The left and right sides of Υ are given by the equations $y = x$ and $y = x - (2 - P)$. Hence Ω maps the left and right sides of Υ respectively to the left and right sides of the integer unit square $[0, 1]^2$. The bottom and top sides of Υ are given by taking

$$c_0 = \frac{P-2}{2}, \quad c_1 = \frac{P-2}{2} + \frac{1-P}{2-P} \quad (220)$$

in Equation 218. We check that $f(c_0) = 1$ and $f(c_1) = 0$. Hence Ω maps the top and bottom of Υ respectively to the bottom and top of $[0, 1]^2$. ♠

The Geometric Claim follows immediately from the analysis in the previous lemma. It says that

$$\Lambda_\Gamma(H^{\text{hi}}) \cap \Upsilon = \Omega^{-1}([0, 1] \times [P, 1]), \quad \Lambda_\Gamma(H^{\text{lo}}) \cap \Upsilon = \Omega^{-1}([0, 1] \times [0, P]). \quad (221)$$

In light of Equation 216, this last equation is equivalent to the Geometric Claim.

This completes the proof of the Geometric Claim, and thereby completes the proof of both the Hitset Theorem and the Intertwining Theorem.

21 Correspondence of Polytopes

In this chapter we prove Statement 2 of the Pixellation Theorem modulo some integer linear algebra calculations. We also set up the kind of problem we need to solve in order to prove Statements 3,4, and 5 of the Pixellation Theorem. We fix some parameter A that is in the background of the whole discussion. Also, as in some previous chapters, we reserve the word *square* for unit integer squares.

21.1 The Triple Partition

Suppose that \mathcal{P}_1 and \mathcal{P}_2 are two partitions of \widehat{X} into polytopes, say

$$\mathcal{P}_k = \bigcup_i P_{i,k}. \quad (222)$$

We define the *common refinement* to be

$$\mathcal{P}_1 \# \mathcal{P}_2 = \bigcup_{i,j} P_{i,1} \cap P_{j,2}. \quad (223)$$

This is the usual definition. In case the polytopes in the two partitions have rational vertices, the polytopes, in the common refinement will also have rational vertices. This construction may be iterated, so that we can take the n -fold common refinement of n partitions.

Let \mathcal{P}_0 be the partition of \widehat{X} that we have described above. We define

$$\mathcal{P}_k = F^k(\mathcal{P}_0). \quad (224)$$

In other words, we take the original partition and apply a power of the PET dynamics. We define the *triple partition* to be the common refinement

$$\mathcal{TP} = \mathcal{P}_{-1} \# \mathcal{P}_0 \# \mathcal{P}_1. \quad (225)$$

We will prove that every vertex of every polytope in \mathcal{TP} has a coordinates which are divisors of 60. That is, if Q is a polytope of \mathcal{TP} , then $60Q$ is an integer polytope.

Each polytope in \mathcal{TP} has a 6 letter label. We simply concatenate the labels for each of the 3 polytopes involved in the intersection, starting with the label for the polytope in \mathcal{P}_{-1} and ending with the label for the polytope

in \mathcal{P}_1 . This 6 letter label has the following meaning. We look at the label of the polytope containing $\Phi_A(c)$ and the label determines the shape of the length 3 arc of the oriented plaid component that is centered at c . Figure 21.1 illustrates this principle with two examples. In Figure 21.1 we show the label and the corresponding path. We hope that these examples suffice to convey the general idea.

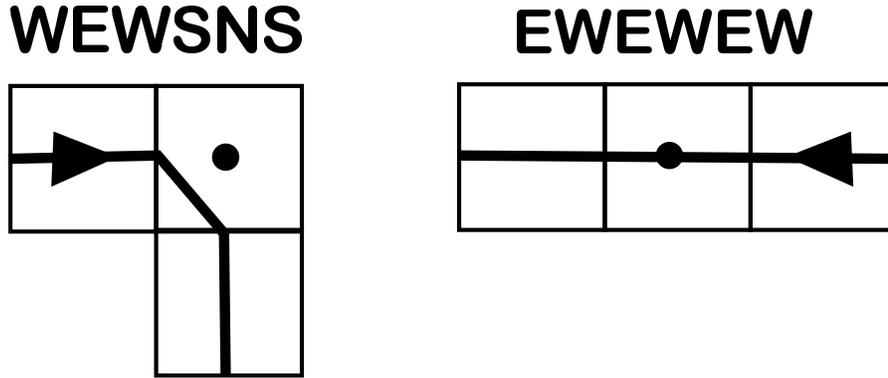


Figure 21.1: The meaning of the 6 letter labels

We define the *reduced triple partition* to be partition

$$\mathcal{RTP} = \mathcal{TP} \#([-1, 1]^3 \times [0, 1]). \quad (226)$$

Essentially we are taking the pieces of \mathcal{TP} which lie inside our favorite fundamental domain for Λ_1 , but in case any of the pieces slop over the boundary of this fundamental domain, we chop them off. Once again, for every $Q \in \mathcal{RTP}$, the scaled polytope $60Q$ is integral.

We think of the polytopes in \mathcal{RTP} as being labeled by the same 6-letter labels, modulo reversal of the labeling. With this interpretation, we can interpret that map $\Phi_A G \rightarrow \widehat{X}$ as a classifying map for unoriented arcs of the plaid model having combinatorial length 3. We just use the action of Λ_1 to move $\Phi_A(c)$ into \mathcal{RTP} and then we read off the label.

There is one minor issue that we need to address. It follows from Theorem 7.5 that $\Phi_A(c)$ always lies in the interior of a polytope of \mathcal{TP} , but it might happen that there are several images of $\Phi_A(c)$ on the boundary of the fundamental domain. This is the usual problem with fundamental domains. However, in this situation, all the images will lie in polytopes having the same labels. So, even when there is some ambiguity in interpreting Φ_A as a map from G into \mathcal{RTP} , the ambiguity is harmless.

Even though it is somewhat less natural than \mathcal{TP} , we will use \mathcal{RTP} for the proof of the Quasi-Isomorphism Theorem because it makes certain formulas simpler.

21.2 Proof of Statement 2

Statement 2 of the Pixellation Theorem says that a grid full square is plaid trivial if and only if it is graph trivial.

The reduced triple partition \mathcal{RTP} consists of 218 polytopes in the fundamental domain $X_{\Pi} = [-1, 1]^3$. We call these polytopes Π_0, \dots, Π_{217} .

Lemma 21.1 *Suppose that Σ is a grid full square. If Σ is plaid trivial then Σ is graph trivial.*

Proof: In our listing, there are 6 polytopes in \mathcal{RTP} having null labels, namely Π_0, \dots, Π_5 . We observe that $\Psi(\Pi_j)$ is contained in a single polytope

$$\Gamma_j = \Gamma_{j,+} = \Gamma_{j,-}$$

in the graph partition, and that Γ_j is null labeled. (This is why $\Gamma_{j,+} = \Gamma_{j,-}$.)

Suppose that Σ is a grid full plaid trivial square for some parameter. Let ζ_{Π} be the center of Σ . Let ζ_{Γ} be the graph grid point in G_{Γ} which is contained in Σ .

By the Plaid Master Picture Theorem,

$$\Phi_{\Pi}(\zeta_{\Pi}) \subset \bigcup_{j=0}^5 \Pi_j.$$

By the Intertwining Theorem,

$$\Phi_{\Gamma}(\zeta_{\Gamma}) \subset \bigcup_{j=0}^5 \Psi(\Pi_j) \subset \bigcup_{j=0}^5 \Gamma_j.$$

By the Graph Master Picture Theorem, the portion of the arithmetic graph associated to ζ_{Γ} is trivial. That is, Σ is graph trivial. ♠

Lemma 21.2 *Suppose that Σ is a grid full square. If Σ is graph trivial then Σ is plaid trivial.*

Proof: We will prove the contrapositive: If Σ is plaid nontrivial then Σ is graph nontrivial. We observe that each of the polytopes Π_6, \dots, Π_{217} has a nontrivial label. We also observe the following:

- $\Psi(\Pi_j)$ is contained in a single graph polytope $\Gamma_{j,+}$ and a single graph polytope $\Gamma_{j,-}$ for $j = 6, \dots, 179$. The labels of these graph polytopes are all nontrivial
- $\Psi(\Pi_j)$ is contained in a single graph polytope $\Gamma_{j,-}$ and a union of two graph polytope $\Gamma_{j+,0}$ and $\Gamma_{j+,1}$ for $j = 180, \dots, 198$. The labels of these graph polytopes are all nontrivial.
- $\Psi(\Pi_j)$ is contained in a single graph polytope $\Gamma_{j,+}$ and a union of two graph polytope $\Gamma_{j-,0}$ and $\Gamma_{j-,1}$ for $j = 199, \dots, 217$. The labels of these graph polytopes are all nontrivial.

Combining the information listed above with the Plaid Master Picture Theorem, the Graph Master Picture Theorem, and the Intertwining Theorem in the same way as the previous proof, we get the conclusion of this lemma. ♠

21.3 A Sample Result

Let us press the method of the previous section a bit harder, in order to see both its power and its limitations. Figure 21.2 shows a particular arc of the plaid model. Depending on the way the triple is oriented, it corresponds to tiles in \mathcal{RTP} labeled NEWEWS or SWEWEN.

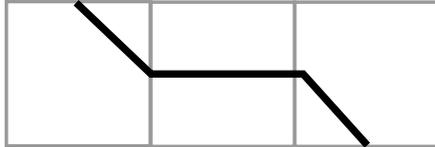


Figure 21.2: The triple of type NEWEWS or SWEWEN.

Lemma 21.3 (Sample) *Every time the triple of the type shown in Figure 21.2 appears in the arithmetic graph and the central square Σ is grid full, the two edges incident to the grid graph vertex in Σ are in $\mathcal{F}(0, 1)$ and $\mathcal{F}(0, -1)$.*

Proof: There are 6 polytopes in \mathcal{RTP} having the relevant labels. The ones labeled NEWNEWS are listed as $\Pi_7, \Pi_{138}, \Pi_{165}$. The ones labeled SWEWEN are listed as $\Pi_{103}, \Pi_{130}, \Pi_{159}$.

Recall that Ψ is the projective intertwining map. We check by direct computation that, for each $j \in \{7, 103, 130, 138, 159, 165\}$ there are polytopes $\Gamma_{j,+}$ and $\Gamma_{j,-}$ in the (+) and (-) graph PET partitions respectively so that

$$\Psi(\Pi_j) \subset \Gamma_{j,+} \cap \Gamma_{j,-}. \quad (227)$$

When we inspect the labels of these graph polytopes we observe that

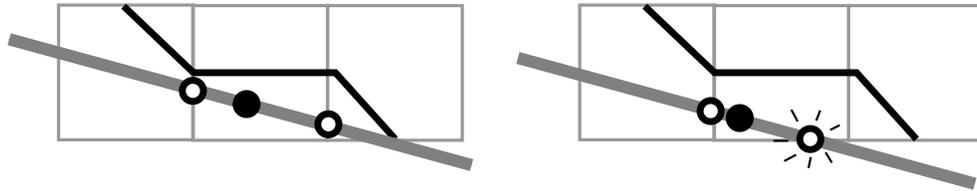
- The label of $\Gamma_{j,+}$ is always $(0, 1)$.
- The label of $\Gamma_{j,-}$ is always $(0, -1)$.

Now let us put this information together. Suppose we see the triple from Figure 21.2 in the plaid model for some parameter A . The rest of the discussion implicitly refers to the parameter A . Let ζ_Γ be the vertex of the graph grid G_Γ contained in the central square Σ of the triple. The same argument as in the previous section shows that

$$\Phi_\Gamma(\zeta_\Gamma) \in \bigcup_{j=7,103,130,138,159,165} \Gamma_{j,+} \cap \Gamma_{j,-}. \quad (228)$$

But then, by the Master Picture Theorem, the two edges incident to ζ_Γ are in $\mathcal{F}(0, 1)$ and $\mathcal{F}(0, -1)$. ♠

10.2 shows us the kind of conclusion we can draw from Lemma 21.3.



10.2: Some possible conclusions from Lemma 21.3.

Lemma 21.3 tells us that every time we see the triple in Figure 21.2, we see the arithmetic graph edges shown in (in grey) in 10.2. However, it might happen that one of these grey edges crosses the top or the bottom of the central square rather than the sides, as the Pixellation Theorem suggests.

Lemma 21.3 tells us the *local geometry* of the arithmetic graph but not how it crosses the edges of the central square. So eliminate the second option shown in 10.2 (as well as other bad options, we need to study how our machinery also determines the edge crossings.

In the rest of this chapter, we set up conventions and notation for the kind of problem we need to solve.

21.4 Fixing Orientations

Recall that there are both oriented and unoriented versions of the Plaid and Graph Master Picture Theorems. We have stated the Intertwining Theorem for the unoriented versions, because the proof is simpler. However, this forces us to deal with orientations in a somewhat *ad hoc* way.

As we mentioned above, there are 218 polytopes in RTP. Each of these polytopes has a 6 letter label which specifies an oriented arc of combinatorial length 3 in the plaid model. The catch is that we are using the small domain \widehat{X}/Λ_1 as the image of the plaid classifying map Φ_Π , rather than the double cover \widehat{X}/Λ_2 .

Thus, if $\zeta_\Pi \in G_\Pi$ is some point, the label of the polytope in \mathcal{RTP} containing $\Phi_\Pi(\zeta_\Pi)$ might specify the opposite orientation on the plaid arc through ζ_Π . More precisely, the assigned orientation is correct if and only if $\widehat{\Phi}_\Pi(\zeta_\Pi)$ is congruent mod Λ_2 to a point in the fundamental domain X_Π . Here $\widehat{\Phi}_\Pi$ is the lift of $\widehat{\Pi}$ to the double cover \widehat{X}/Λ_2 . At the same time, in order to figure out the orientations on the graph polygons, we would need to look at the lift $\widehat{\Phi}_\Gamma$ of the graph classifying map Φ_Γ .

Getting these orientations right is a tedious business, and we have a different way of dealing with it. We ignore the “true” orientations coming from the lifted maps and we simply make a guess as to how the plaid arcs correspond to the graph arcs. For example, in 10.2, it is pretty clear that the grey edge pointing left should be associated to the left half of the plaid triple.

Now we explain the convention in more detail, by way of example. Figure 21.3 shows an enhanced version of 21.2, specifically for Π_7 . Again, the label of Π_7 is NEWNEWS, and this determines the orientation.

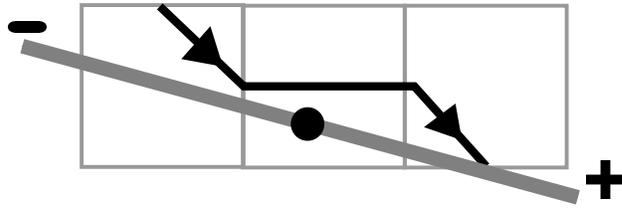


Figure 21.3: The case of Π_7 : a correspondence of type 1

We observe experimentally that the $(-)$ partition assigns the label $(0, -1)$, and this corresponds to the leftward pointing edge. Likewise, the $(+)$ partition assigns the label $(0, 1)$, and this corresponds to the rightward pointing edge. Thus, we associate the $(-)$ partition with the tail end of the triple and the $(+)$ partition with the head end. For this reason, we call Π_7 *type 1*. We would call Π_j *type 0* if, according to our experimental observations, the $(+)$ edge is associated with the tail and the $(-)$ edge is associated with the tail.

We guess the type for each of the 218 polytopes in the plaid triple partition, and these types are stored in the computer program. In most cases, as for Π_7 , the picture is completely obvious, and in a few cases, like the one shown in 10.2 for Π_{49} the picture is not quite as obvious but still pretty clear.

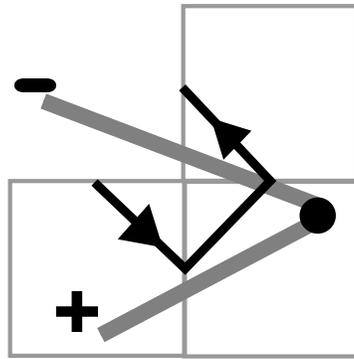


Figure 21.4: The case of Π_{48} : a correspondence of type 0.

Logically, it does not matter how we arrived at these guesses, or whether they agree with the true answer which can be cleaned by looking at the lifts. The point is simply that the proof runs to completion with the guesses made. In hindsight, our guesses surely agree with the true answer, but we do not need to prove this and we will not.

21.5 Edge Crossing Problems

To each index $i = 6, \dots, 218$ we associate what we call an *edge crossing problem*. We will first give the general definition and then we will work out the example of Π_7 .

Given Π_k , there corresponds a triple of squares $\Sigma_{-1}, \Sigma_1, \Sigma_1$ and an oriented plaid arc running through it, as in Figures 21.3 and 21.4. We label the squares so that Σ_1 is the head square in the type 1 case and the tail square in the type 0 case. In other words, the arithmetic graph edge associated to the (\pm) partition should point generally from Σ_0 to Σ_{\pm} for both types.

There are two kinds of edge crossing problems associated to P_k . One kind is labeled $(k, +, i, j, L)$. Here $L \in \{N, S, E, W\}$ is an edge of Σ_0 that the arithmetic graph edge $dT(i, j)$ associated to the $(+)$ partition could potentially cross (as in the right hand side of Figure 21.3) but according to the Pixellation Theorem is not supposed to cross. These crossing problems really only depend on the pair (Σ, Σ_+) . The other kind of edge crossing problem is labeled $(k, -, i, j, L)$, and has a similar explanation with $(-)$ in place of $(+)$.

Let's consider the case of P_7 . The two crossing problems are $(7, +, 0, 1, N)$ and $(7, -, 0, -1, S)$. The other two possibilities, namely $(7, +, 0, 1, E)$ and $(7, -, 0, -1, W)$, are the ones predicted by the Pixellation Theorem. They are not problems at all, but rather *goals*.

We have already mentioned that the first 6 of the plaid triple polytopes correspond to the trivial grid full squares. There are another 174 polytopes, labeled Π_6, \dots, Π_{179} , which have the property that there are unique graph polytopes $\Gamma_{k,+}$ and $\Gamma_{k,-}$ such that

$$\Psi(\Pi_k) \subset \Gamma_{k,\pm}. \quad (229)$$

Each of these contributes 2 edge crossing problems, giving a total of 348.

The 19 polytopes $\Pi_{180}, \dots, \Pi_{198}$ are such that

$$\Psi(\Pi_k) \subset \Gamma_{k,+,0} \cup \Gamma_{k,+,1}, \quad \Psi(\Pi_k) \subset \Gamma_{k,-}. \quad (230)$$

That is, $\Psi(\Pi_k)$ is contained in a union of two graph polytopes from the $(+)$ partition and 1 from the $(-)$ partition. In this case, there are two possible local picture of the arithmetic graph associated to this plaid triple. This does not bother us, as long as the edge crossings come out right. Each of these 19 polytopes contributes 3 edge crossing problems. This gives us another 57.

The 19 polytopes $\Pi_{199}, \dots, \Pi_{217}$ are such that

$$\Psi(\Pi_k) \subset \Gamma_{k,-,0} \cup \Gamma_{k,-,1}, \quad \Psi(\Pi_k) \subset \Gamma_{k,+}. \quad (231)$$

Each of these 19 polytopes contributes 3 edge crossing problems, giving yet another 57 crossing problems.

The grand total is 462 edge crossing problems. In the next chapter we will introduce the machinery needed to solve all these problems, so to speak. A *solution* amounts to a proof that the given case does not actually occur. Actually, we will be able to solve 416 of the problems. The remaining 46, in the cases where the pixellation really does fail, are involved in the catches for the offending edges discussed in the Pixellation Theorem. Once we have solved all the problems we can solve, and classified the exceptions, the rest of the Pixellation Theorem just comes down to inspecting the data generated by the program.

22 Edge Crossings

22.1 The Graph Method

Here we explain the first method we use for solving the edge crossing problems. We fix a parameter A throughout the discussion. Let G_Γ denote the grid graph. Let $\zeta_\Gamma \in G$ be some point, contained in a square Σ . As in previous chapters, the word *square* always means a unit integer square.

We assume that ζ_Γ is a nontrivial vertex of the arithmetic graph. Let e be one of the edges of the arithmetic graph incident to ζ_Γ . We think of e as a vector pointing from v out of Σ . From the Grid Geometry Lemma, we know that e crosses some edge of Σ . We say that e is of type (i, j, L) if $e = dT(i, j)$ and $L \in \{S, W, N, E\}$ is the label of the edge of Σ which e crosses.

Remark: We allow the possibility that e is of two types. This would happen if e crosses Σ at a vertex. We think that this never actually happens, but we have not ruled it out. In any case, this eventuality does not bother us.

Let Φ_Γ denote the graph classifying map. We define

$$(x, y, z, A) = \Phi_\Gamma(\zeta_\Gamma) \tag{232}$$

Lemma 22.1 (Graph Avoidance) *The following is true.*

1. If $x \in (A, 1)$ then e is not of type $(-1, 0, W)$.
2. If $x \in (0, A)$ then e is not of type $(-1, 1, W)$.
3. If $y \in (A, 1)$ then e is not of type $(0, -1, N)$.
4. If $y \in (2A, 1 + A)$ then e is not of type $(0, 1, S)$.
5. If $x \in (1 - A, 1)$ then e is not of type $(1, -1, E)$.
6. If $x \in (0, 1 - A)$ then e is not of type $(1, 0, E)$.

Proof: We will treat the first three cases. The last three cases follow from symmetry. More precisely, reflection in the center of the square $[0, 1]^2$ carries the sets used to analyze Case N to the sets used to analyze Case 7-N. Figure 22.1 illustrates our arguments.

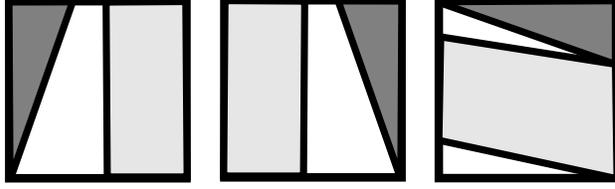


Figure 22.1: The first three cases

Case 1: We have

$$dT(-1, 0) = \left(-A, -\frac{1 + 2A - A^2}{1 + A} \right).$$

This vector points southwest and has slope greater than 1. So, in order for this edge to cross the west edge, it must lie in the triangle with vertices

$$(0, 0), \quad (0, 1), \quad \left(\frac{A^2 + A}{1 + 2A - A^2} \right).$$

We call this triangle *the danger zone*. The line through this third point and parallel to e contains the vertex between the south and west edges. The dark triangles in Figure 22.1 are hand-drawn versions of the danger zones in each case.

The Graph Reconstruction Formula tells us that the conditions $x \in (A, 1)$ correspond to the condition that the first coordinate of ξ_{Γ} lies in $(A, 1)$. The region of possibilities is the lightly shaded region on the left hand side of figure 23.1. As depicted in the figure, the two sets we have defined are disjoint.

Case 2: The argument is the same, except this time the danger zone has vertices

$$(1, 0), \quad (1, 1), \quad \left(\frac{4A}{4A - A^2} \right),$$

and the condition $x \in (0, A)$ corresponds to these same conditions on the first coordinate of ξ_{Γ} . Again, the two sets are disjoint. This is shown in the middle square of Figure 22.1.

Case 3: The argument is the same, except this time the danger zone has vertices

$$(0, 1), \quad (1, 1), \quad \left(1, \frac{1 - A}{1 + A} \right).$$

and the condition $y \in (A, 1)$ corresponds, *via* the Reconstruction Formula to ξ_Γ lying in the interior of the parallelogram with vertices

$$\left(0, \frac{A}{1+A}\right), \quad \left(0, \frac{1}{1+A}\right), \quad \left(1, \frac{1-A}{1+A}\right), \quad (0, 0).$$

Again, these sets are disjoint. This case is shown on the right side of Figure 22.1. ♠

Remark: In the previous result, the cases $(1, 1, N)$ and $(-1, -1, S)$ are missing. There is a similar result for these cases, but we will prove a result below that is more powerful and subsumes these cases. So, we ignore them here.

22.2 The Sample Result Revisited

In this section we explain how we solve the crossing problem $(7, +, , 0, 1, S)$. In other words, we are trying to rule out the bad crossing indicated on the right hand side of Figure 21.2, which we repeat here with enhanced labeling. We really need to solve 6 crossing

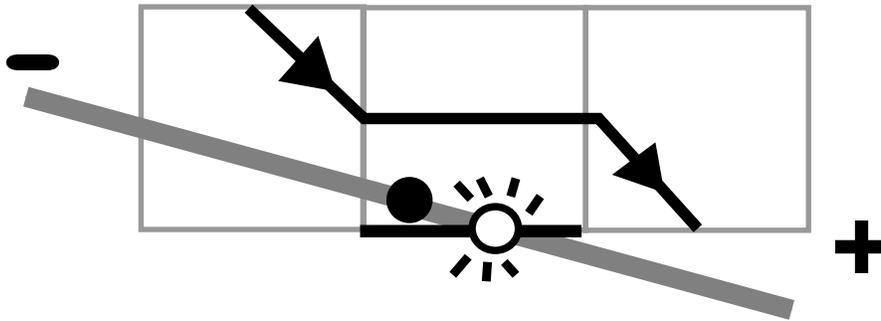


Figure 22.2: The crossing problem $(7, +, S)$.

The polytope $\Psi(\Pi_7)$ has 8 vertices:

- $(60, 60, 0, 30)/60$.
- $(60, 40, 20, 20)/60$.
- $(40, 40, 0, 20)/60$.
- $(60, 40, 0, 20,)/60$.

- $(60, 60, 0, 20)/60$.
- $(45, 45, 0, 15)/60$.
- $(60, 45, 15, 15)/60$.
- $(60, 45, 0, 15)/60$.

We have written the vertices this way so as to clear denominators. The factor of 60 works for every polytope in sight. (My program has a window which allows the user to see the vertices of any polytope in any of the partitions.)

Recalling that we coordinatize \widehat{X} using the variables, note that

$$y \in [2A, 1 + 2A]. \quad (233)$$

for all vertices of $\Psi(\Pi_7)$. By convexity, this equation holds for all points, and we get strict inequality for points in the interior of $\Psi(\Pi_7)$.

It follows from the Master Picture Theorems and the Intertwining Theorem that

$$(x, y, z, A) = \Phi_\Gamma(\zeta_\Gamma) \in \text{interior}(\Psi(\Pi_y)). \quad (234)$$

Hence $y \in (2A, 1 + 2A)$. Case 4 of the Graph Avoidance Lemma, which pertains to the triple $(0, 1, S)$, solves this crossing problem.

The argument works the same way for all 6 crossing problems associated to the bad crossing shown in Figure 22.2. It follows from symmetry (or from a similar argument) that the 6 crossings associated to the other arithmetic graph edge in Figure 22.2 are also soluble. Thus, every time this pattern occurs in the plaid model for any parameter, and the central square is grid full, the central square is pixellated.

Remark: Sometimes we use the Graph Avoidance Lemma in a different way. To illustrate our other usage, we will solve the edge crossing problem $(198, +, -1, 1, E)$. This time it turns out that the criterion in the Graph Avoidance Lemma for $(1, -1, E)$ does not hold for $\Psi(\Pi_{198})$. However, it does hold for $\Gamma_{198,+0}$, the relevant one of the two graph polytopes in the $(+)$ graph partition whose union contains $\Psi(\Pi_{198})$. So, again, we find that $\Phi_\Gamma(\xi_\Gamma)$ cannot lie in the region corresponding to a situation where the graph edge incident to ζ_Γ crosses E .

For all the relevant crossing problems on which we use the Graph Avoidance Lemma, we will either apply the criteria to $\Psi(\Pi)$ or to the relevant graph polytope Γ . In short, we will use the one method or the other.

22.3 The Plaid Method

Here we discuss a second method for solving crossing problems. In the discussion that follows, $L \in \{N, S, E, W\}$ stands for one of the edge labels. To 10 of the 16 pairs (i, j, L) . Let $\langle i, j, L \rangle$ denote the subset of \widehat{X} which assigns the edge $e = dT(i, j)$ to grid graph points ζ_Γ . in such a way that e crosses the edge L of the square containing the point ζ_Γ . The Graph Avoidance Lemma can be interpreted as saying that certain regions in \widehat{X} avoid $\langle i, j, L \rangle$. For instance, the set $x \in (A, 1)$ is disjoint from $\langle -1, 0, W \rangle$.

Recall that X_Π^* is the hitset. We are going to associate a polytope $Z(i, j, L) \subset X_\Pi$ such that one of two things is true:

- If $P \cap Z(i, j, L) = \emptyset$ then $\Psi(P) \cap \langle i, j, L \rangle = \emptyset$.
- If $P \subset Z(i, j, L)$ then $\Psi(P \cap X_\Pi^*) \cap \langle i, j, L \rangle = \emptyset$.

In the first case, we call P an *excluder* and in the second case we call P a *confiner*. The other 6 pairs we simply ignore.

For each fixed parameter, our polytopes all have the form

$$Z(i, j, L) = Z'(i, j, L) \times [-1, 1], \quad (235)$$

Where $Z'(i, j, L)$ is a polygon in the xy plane. We will list 4 of the 10 sets. The other 5 are obtained from the first 5 *via* the following symmetry.

$$Z(-i, -j, L^{\text{opp}}) = -Z(i, j, L). \quad (236)$$

Here L^{opp} is defined to be the edge opposite L . For instance $E^{\text{opp}} = W$. We list the 5 polygons $Z'(i, j, L)$ as a function of the parameter A , and recall that $P = 2A/(1 + A)$. The first set is an excluder and the other 5 are confiners. Here are the sets:

- $Z'(1, 1, N)$: $(1 - P, 1 - P), (P - 1, P - 1), (P - 1, -1), (1 - P, -1)$.
- $Z'(1, 1, E)$: $(1 - P, 1 - P), (-1, -1), (1 - P, -1)$.
- $Z'(0, 1, S)$: $(-1, -1), (P - 1, P - 1), (1, P - 1), (1 - P, -1)$.
- $Z'(1, 0, E) = Z'(1, 1, E)$.
- $Z'(0, -1, W) = Z'(0, 1, S)$.

Conveniently, all these sets intersect the hitset inside the fundamental dipole Υ considered in §20.10. This makes our analysis easy.

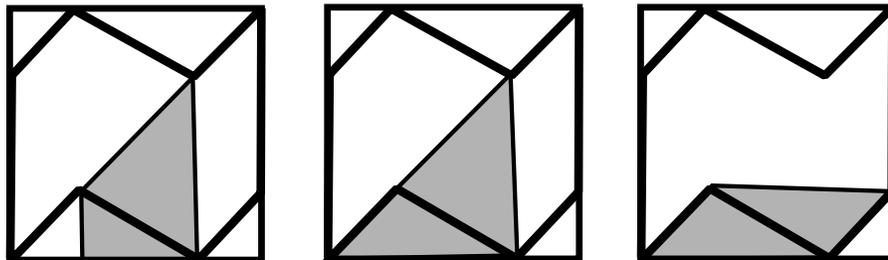


Figure 22.4: The barrier bases in action.

To establish our claims, we use the Plaid Reconstruction Formula, Equation 216, to map our sets into the unit square, and we check that the relevant image is disjoint from the set of positions in $[0, 1]^2$ where ζ_Γ can be placed so that $dT(i, j)$ crosses edge L . We called these sets the danger zones in the proof of the Graph Avoidance Lemma.

22.3.1 Case 1

Here we consider $(1, 1, N)$. Calculations like the one done in the proof of the Graph Avoidance Lemma show that the danger zone is the triangle with vertices

$$(1, 1), \quad (0, 1), \quad (0, P).$$

Let $Z'_*(1, 1, N)$ denote the intersection of $Z'(1, 1, N)$ with the planar projection of the hitset. The vertices of $Z'_*(1, 1, N)$ are

$$(P - 1, P - 1), \quad (1 - P, -1), \quad (1 - P, 1 - P).$$

Beautifully, the image of this triangle under the affine diffeomorphism from Equation 216 is exactly the danger zone.

22.3.2 Case 2

Here we consider $(1, 1, E)$. The danger zone is the complement of the triangle considered in Case 1. We have $Z'_+(1, 1, N) = Z'_+(1, 1, E)$. So, we have already computed the relevant affine image; it is the set $\Sigma_0(1, 1)$ from Case 1. But the interiors of $\Sigma_0(1, 1)$ and $\Sigma_1(1, 1)$ are disjoint: these two sets partition the unit square.

22.3.3 Case 3

Here we consider $(0, 1, S)$. The danger zone is the same as the one in Case 4 of the Graph Avoidance Lemma. It has vertices

$$(0, 0), \quad (1, 0), \quad (0, P).$$

The set $Z'_*(0, 1, Z)$ is the triangle with vertices

$$(P - 1, P - 1), \quad (1 - P, -1), \quad (1, P - 1).$$

The image of this set under the affine diffeomorphism is the triangle with vertices

$$(0, 1), \quad (1, 1), \quad (1, P).$$

This triangle is clearly disjoint from the danger zone.

22.3.4 Case 4

Here we consider $(1, 0, E)$. The danger zone has vertices

$$(1, 1), \quad (1, 0), \quad \left(\frac{1 + A - 2A^2}{1 + 2A - A^2} \right)$$

We have $Z'(1, 0, E) = Z'(1, 1, E)$. The analysis in Case 1 shows that the affine image of $Z'_+(1, 0, E)$ is the triangle with vertices

$$(1, 1), \quad (0, 1), \quad (0, P).$$

The interior of this set is clearly disjoint from the danger zone.

22.3.5 Case 5

Here we consider $(0, -1, W)$. The danger zone has vertices

$$(0, 0), \quad (1, 0), \quad (1, P).$$

We have $Z'(0, -1, W) = Z'(0, 1, S)$, the set from Case 3. As in Case 3, the vertices of the affine image of $Z'(0, -1, W)$ is the triangle with vertices

$$(0, 1), \quad (1, 1), \quad (1, P).$$

The interior of this set is disjoint from the danger zone.

22.4 Another Edge Crossing Problem

Here we solve another edge crossing problem, to illustrate the logic behind the use of the Plaid Method. We use the notation from above. One of the crossing problems is $(50, +, -1, -1, W)$. This time we find that $\Pi_{50} \subset Z(-1, -1, W)$. Therefore Hence

$$\Psi_{\Pi}(\zeta_{\Pi}) \subset Z(-1, -1, W). \quad (237)$$

The existence of ζ_{Γ} means that the relevant square Σ is grid full.

$$\Psi_{\Pi}(\zeta_{\Pi}) \subset Z(-1, -1, W) \cap X_{\Pi}^*. \quad (238)$$

By the Intertwining Theorem

$$\Phi_{\Gamma}(\zeta_{\Gamma}) \subset \Psi(Z(-1, -1, W) \cap X_{\Pi}^*). \quad (239)$$

By the confining property of $Z(-1, -1, W)$,

$$\Phi_{\Gamma}(\zeta_{\Gamma}) \notin \langle -1, -1, W \rangle. \quad (240)$$

This solves the crossing problem.

22.5 Out of Bounds

Sometimes none of the above methods works for a crossing problem, but then we notice that Π_k lies entirely outside the hitset Z_{Π}^* . In this case we call Π_k *out of bounds*. The corresponding crossing problem simply does not arise for a grid full square.

We will verify that Π_k is out of bounds by showing that

$$\Pi_k \subset \bigcup_{i=1}^4 B_i \times [-1, 1]. \quad (241)$$

Where B'_1, B'_2, B'_3, B'_4 are the following 4 triangles, described in terms of their vertices:

1. $B'_1 : (-1, -1), (P - 1, -1), (P - 1, P - 1)$.
2. $B'_2 : (1 - P, -1), (1, -1), (1, P - 1)$.
3. $B'_3 = -B'_1$.

4. $B'_4 = -B'_2$.

Figure 22.5 shows how these triangles sit in relation to the planar projection of the hitset.

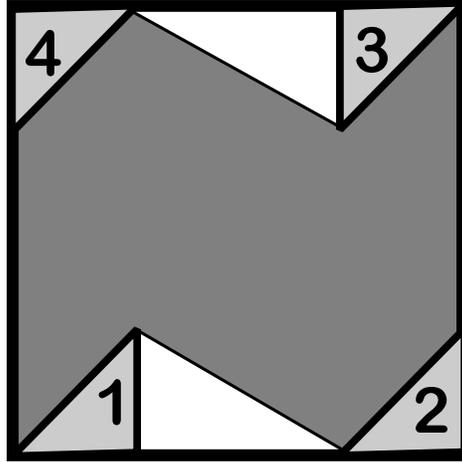


Figure 22.5: The out-of-bounds polygons in action.

Note that the union

$$B_i = \bigcup_{P \in [0,1]} (B'_i \times \{P\}) \quad (242)$$

is a convex polytope with integer coefficients. Here $P = 2A/(1+A)$ as usual. This, showing that $\Pi_k \subset B_i$ for some pair (k, i) is just a matter of integer linear algebra. We call the verification that $\Pi_k \subset B_i$ the *out of bounds test*.

23 Proof of the Pixellation Theorem

Pixellation Theorem. In this chapter, we prove Statements 3,4,5 of the Pixellation Theorem modulo certain integer computer calculations.

23.1 Solving Most of the Crossing Problems

Using the methods discussed in the last chapter, namely the Graph Method, the Plaid Method, and the Out of Bounds Test, we solve 416 of the 462 crossing problems. Each case is like one of the ones considered in the previous chapter, but there are too many to do by hand. In the next chapter we will explain the integer linear algebra tests we use to check each case.

There are 46 exceptional cases. Let's call these 46 cases *recalcitrant*. A few of the recalcitrant cases actually are soluble by more delicate methods - we would just need to look harder at the polytopes involved - but there isn't any problem in our proof with leaving them unsolved.

23.2 Proof of Statement 3

Figure 17.3 shows examples of an errant edge when the two squares are stacked side by side. We repeat the picture here. We will consider the picture on the right hand side of Figure 23.1 in detail. In this picture, the square with the dot is the central square and the top square is not part of the triple. So, we are only showing two of the three squares in the triple.

The codes associated to the figure on the right are either LLLEWS or SWELLL, depending on the orientation. Here $L \in \{N, S, E, W\}$ is a label we don't know. The corresponding errant edge must rise up at least 1 unit. This leaves $dT(1, -1)$ and $dT(1, 0)$ as the only possibilities.

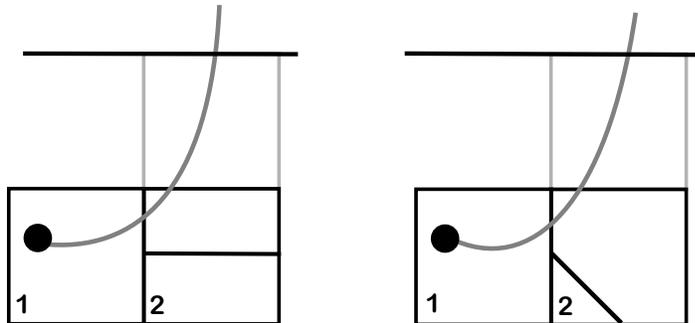


Figure 23.1: errant edges

From the correspondence of polytopes explained in §21, we know the complete list of labels associated to relevant ends of the relevant triples. We just check, by a quick computer search (and also by inspection) that the errant labels never arise. Thus, for instance, we never see the label $(1,0)$ associated to the head end of LLEWS or the tail end of SWELLL. We rule out the other possibilities similarly. Here is a list of the possible codes and the forbidden labels. To save space, we only list half the possibilities. The other half are obtained from these by 180 degree rotation. Also, of the half we do list, we only list the codes such that the corresponding errant edge would be associated to the head of the triple. So, we would list LLEWS only in our example above. Here is the list.

- LLEWS: $(1,0)$, $(1,-1)$.
- LLEWN: $(-1,0)$, $(-1,1)$.
- LLEWE: $(1,0)$, $(1,-1)$, $(-1,0)$, $(-1,1)$.
- LLLSNE: $(-1,-1)$, $(0,-1)$.
- LLLSNW: $(1,1)$, $(0,1)$.
- LLLSNS: $(-1,-1)$, $(1,1)$, $(0,1)$, $(0,-1)$.

We simply check that none of these bad situations actually arises. The rest of the proof of the Pixellation Theorem is devoted to proving Statements 3 and 4. This takes more work.

23.3 Proof of Statement 4

Statement 4 of the Pixellation Theorem says that two arithmetic graph edges incident to a vertex in a square never cross the same edge of that square. This result follows immediately for the 416 cases in which can solve the crossing problem: These cases are all pixellated and the edges in question cross the same sides as the plaid model segments. It just remains to deal with the 46 recalcitrant cases.

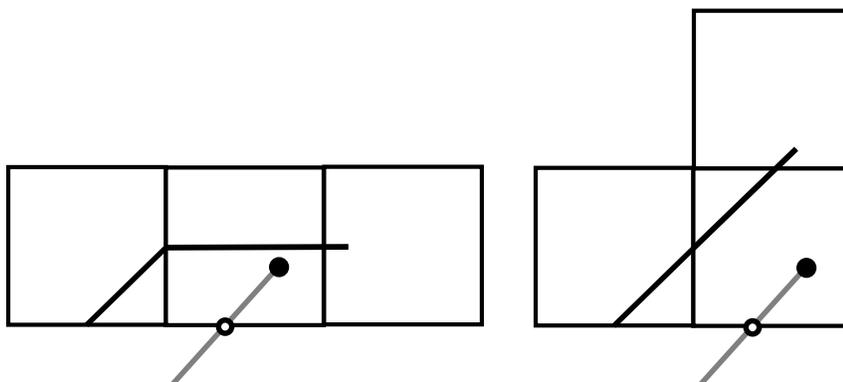


Figure 23.2: The recalitrant cases

By direct inspection, we see that all of the recalitrant cases are either pixellated or equivalent to the two cases shown in Figure 23.2. (By symmetry, we just have to inspect 23 of the 46 cases.) By *equivalent* we mean that the picture is meant to be taken up to rotations and reflections.

Now consider the arithmetic graph edges associated to the a recalitrant case. The first edge, the one shown in Figure 23.2, is offending. The other edge is either offending or not. If the other edge is not offending, then it manifestly crosses a different edge of the central square. If the other edge is offending, and crosses the same edge as the first offending edge, then the picture must look like Figure 23.3.

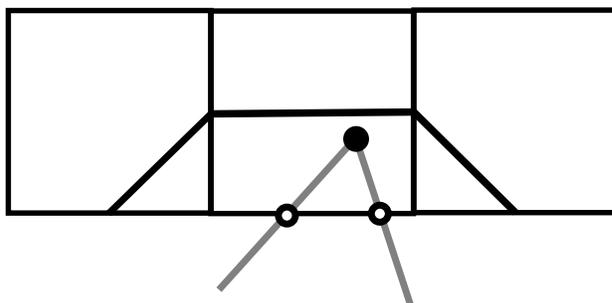


Figure 23.3: A double edge crossing

There are exactly two cases like this, corresponding to P_{34} and P_{173} . But, in both cases, we check that the polytope is involved in only one recalitrant crossing problem. So, in these two cases, only one of the two edges is offending.

This takes care of all the possibilities.

23.4 Proof of Statement 5

Each recalcitrant case corresponds to a plaid triple. Using the curve-following dynamics discussed in §7.5, we check which plaid triples could attach to the one we have. We will illustrate what we mean by example and then state the general result.

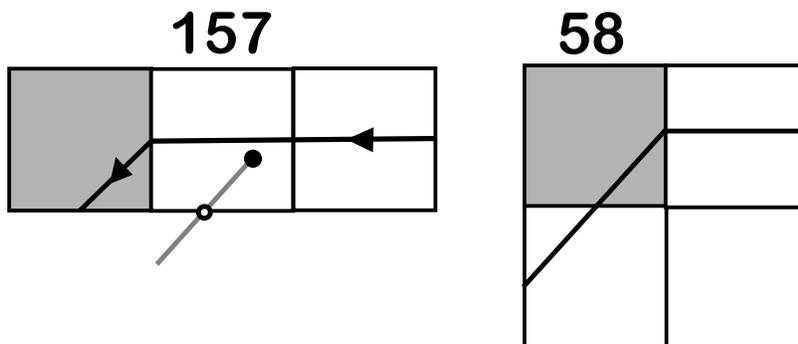


Figure 23.4: The triples corresponding to Π_{157} and Π_{58} .

One of the recalcitrant triples is $(157, -, -1, 0, S)$. The associated code for Π_{157} is EWEWES. Figure 12.4 shows the situation. The offending edge goes with the orientation of the plaid arc, and so we do the $(+)$ dynamics to figure out what happens to points in Π_{157} . That is, we consider the image $F(\Pi_{157})$. (In the other case, when the offending edge goes against the orientation, we would use the map F^{-1} .) We prove that

$$F(\Pi_{157}) \subset \Lambda(\Pi_{58}). \quad (243)$$

Here Λ is the plaid lattice.

We then check that Π_{58} is grid empty. This means that any unit integer square in the plaid model classified by Π_{58} is grid empty. Here we mean that Φ_{Π} maps the center of the square into Π_{58} . But then Equation 243 tells us that the square on the right and side of any plaid triple associated to Π_{157} is grid empty. We have shaded the grid empty squares.

Equation 243 says that every occurrence of a plaid triple associated to Π_{157} is conjoined, so speak, with a plaid triple associated to Π_{58} . We have also added in an extra square, even though we don't know the picture in this square. Figure 12.5 shows the situation.

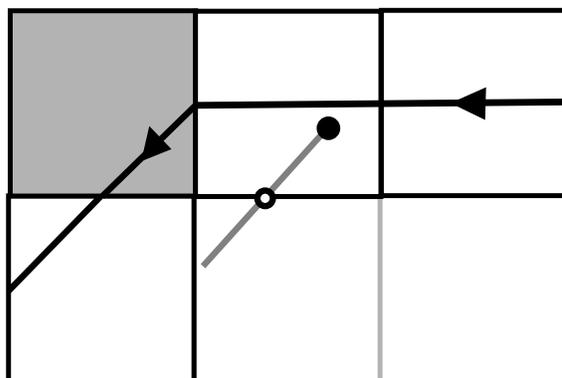


Figure 23.5: The plaid quadruple obtained by concatenating Π_{157} and Π_{38} .

When we look at the curve following dynamics for the other recalcitrant cases, we discover that the same thing always happens: The side square associated to the offending edge is grid empty. Up to isometry, the picture always looks like one of the cases of Figure 23.6.

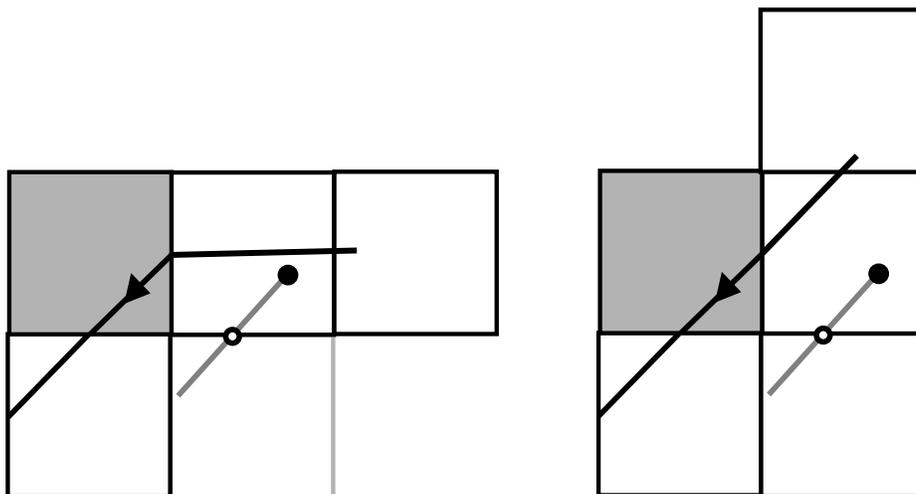


Figure 23.6: Concatenations for the recalcitrant triples

The only thing we need to do in order to finish the proof of Statements 3 and 4 is to analyze where the offending edge ends. There are 8 cases to consider, and the last 4 cases are rotated images of the first 4 cases. So, we will just consider the first 4 cases.

23.4.1 Case 1

In this case, the offending edge is $dT(-1, 0)$ and the picture is oriented as in Figure 23.6. There are 9 cases like this. We have

$$dT(0, -1) = \left(-A, \frac{A^2 - 2A - 1}{1 + A} \right).$$

For each of the 9 cases, we check that

$$\Psi(\Pi_k) \subset \{(x, y, z, A) | x \leq A\}. \quad (244)$$

By the Reconstruction Formula, the first coordinate of $(a, b) = \zeta_\Gamma$, the graph grid point contained in the relevant square, to satisfy $[a] < A$. But then the offending edge must cross over the thick vertical line shown in Figure 23.7 and end in one of the two indicated squares.

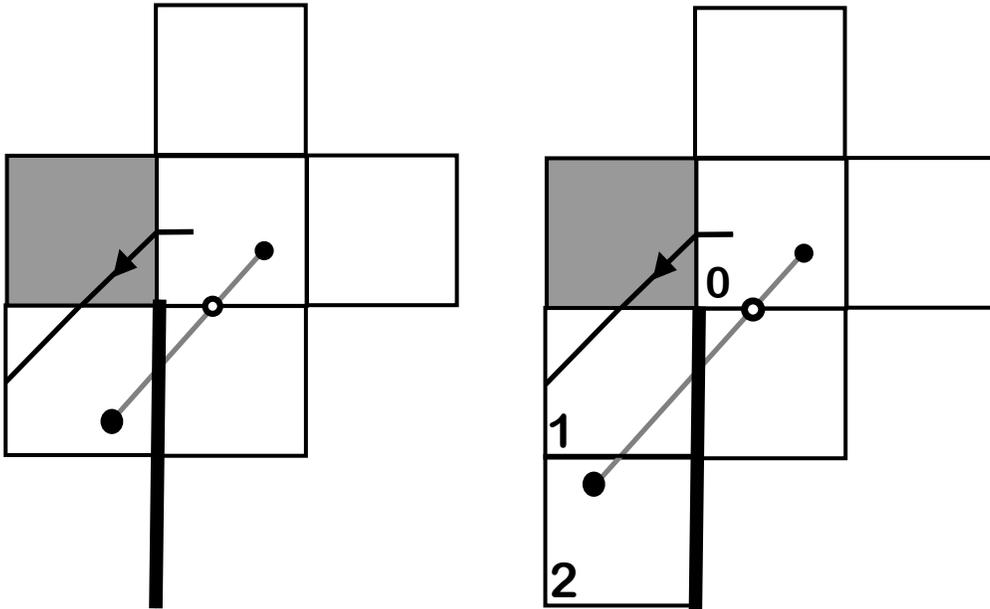


Figure 23.7: The two possible endings for the offending edge

We recognize the two cases as reflected versions of the catches in Figure 4.2. The proof is done in this case.

23.4.2 Case 2

In this case, the offending edge is $dT(-1, 0)$ and the picture is oriented as in Figure 23.8. There are 7 cases like this, and they all have the features shown in Figure 23.8.

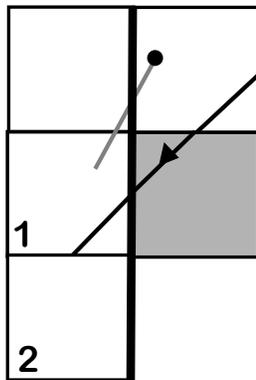


Figure 23.8: Cast 2

We check that Equation 244 holds in all 7 cases. Hence, the offending edge crosses the thick vertical line and ends in the squares marked 1 and 2. Again, we have the catches shown in Figure 4.2.

23.4.3 Case 3

In this case, the offending edge is $dT(0, -1) = (-1, P)$, and the picture is oriented as in Figure 23.9. There are 3 cases like this.

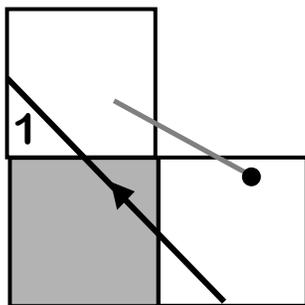


Figure 23.9: Case 3

This case is easy. We know that the offending edge crossed the top of the square it starts in, and given that the vector is $(-1, P)$, it must end in the square Σ_1 . This gives us the left hand catch in Figure 4.2, up to orientation.

23.4.4 Case 4

In this case, the offending edge is $dT(0, -1) = (-1, P)$, and the picture is oriented as in Figure 23.10. There are 4 cases like this.

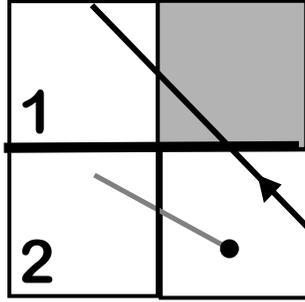


Figure 23.10: Case 4

In this case, the offending edge crosses the thick horizontal line provided that the coordinates (a, b) of ζ_Γ satisfy

$$b + Pa \geq 1 - P, \quad b \leq 1. \quad (245)$$

Notice that this is a weaker condition than $b + Pa \geq 1$, which is what we would have needed to solve the crossing problem. Using the Graph Reconstruction Formula, we have

$$a = x, \quad b = \frac{y - Ax}{1 + A} \quad (246)$$

Plugging this equation into Equation 245 and simplifying, we find that Equation 245 holds provided that

$$y \in [1 - A - Ax, 1 + A + Ax] \quad (247)$$

We check, for each of the cases, that $x \geq 0$ and $y \in [1 - A, 1 + A]$. This does the job for us. Thus, the offending edge ends in the square Σ_1 , and we get the catch on the left hand side of Figure 4.2.

In all cases, each offending edge has a catch. At the same time, each catch that appear

24 Computer Assisted Techniques

24.1 Operations on Polytopes

Clean Polytopes: Say that a *clean polytope* is a convex polytope in \mathbf{R}^4 with integer vertices, such that each vertex is the unique extreme point of some linear functional. In other words, a clean polytope is the convex hull of its (integer) vertices, and the convex hull of any proper subset of vertices is a proper subset. We always deal with clean polytopes. The polytopes in the plaid and graph triple partitions have all the properties mentioned above, except that their vertices are rational rather than integral. We fix this problem by scaling all polytopes in all partitions by a factor of 60.

Clean Polytope Test: Suppose we are given a finite number of integer points in \mathbf{R}^4 . Here is how we test that they are the vertices of a clean polytope. We consider all linear functionals of the form

$$L(x, y, z, A) = c_1x + c_2y + c_3z + c_4A, \quad |c_i| \leq N \quad (248)$$

and we wait until we have shown that each vertex is the unique maximum for one of the functionals. For the polytopes of interest to us, it suffices to take $N = 3$. In general, our test halts with success for some N if and only if the polytope is clean.

Disjointness Test: Here is how we verify that two clean polytopes P_1 and P_2 have disjoint interiors. We consider the same linear functionals as listed in Equation 248 and we try to find some such L with the property that

$$\max_{v \in V(P_1)} L(v) \leq \min_{v \in V(P_2)} L(v). \quad (249)$$

Here $V(P_k)$ denotes the vertex set of P_k . If this happens, then we have found a hyperplane which separates the one polytope from the other. This time we take $N = 5$.

Containment Test: Given clean polytopes P_1 and P_2 , here is how we verify that $P_1 \subset P_2$. By convexity, it suffices to prove that $v \in P_2$ for each vertex of P_1 . So, we explain how we verify that $P = P_2$ contains an integer point v . We do not have the explicit facet structure of P , though for another purposes (computing volumes) we do find it.

Let $\{L_k\}$ denote the set of all linear functionals determined by 4-tuples of vertices of P . Precisely, Given 4 vertices of P , say w_0, w_1, w_2, w_3 , and some integer point v , we take the 4×4 matrix whose first three rows are $w_i - w_0$ for $i = 1, 2, 3$ and whose last row is v . Then

$$L_{w_0, w_1, w_2, w_3}(v) = \det(M) \quad (250)$$

is the linear functional we have in mind.

If the vertices do not span a 3-dimensional space, then L will be trivial. This does not bother us. Also, some choices of vertices will not lead to linear functions which define a face of P . This does not bother us either. The point is that our list of linear functionals contains all the ones which do in fact define faces of P . We take our vertices in all orders, to make sure that we pick up every possible relevant linear functional. The computer does not mind this redundancy.

It is an elementary exercise to show that $v \notin P$ if and only if v is a unique extreme point amongst the set $\{v\} \cup V(P)$ for one of our linear functionals. Here $V(P)$ denotes the vertex set of P , as above. So, $v \in P$ if and only if v is never a unique extreme point for any of the linear functionals on our list.

Volume: First we explain how we find the codimension 1 faces of P . We search for k -tuples of vertices which are simultaneously in general position and the common extreme points for one of the linear functionals on our list. As long as $k \geq 4$, the list we find will be the vertices of one of the faces of P .

Now we explain how we compute the volume of P . This is a recursive problem. Let v_0 be the first vertex of P . Let F_1, \dots, F_k be the codimension 1 faces of P . Let V_j be the cone of F_j to v_0 . This is the same as the convex hull of $F_j \cup \{v_0\}$. Then

$$\text{vol}(V) = \sum_{j=1}^k \text{vol}(V_j). \quad (251)$$

If $v_0 \in V_j$ then the volume is 0. These extra trivial sums do not bother us.

To compute $\text{vol}(V_j)$ we let $w_{j0}, w_{j1}, w_{j2}, w_{j3}$ be the first 4 vertices of F_j . Let L_j be the associated linear functional. Then

$$4 \times \text{vol}(V_j) = L_j(v_0 - w_{j0}) \times \text{vol}(F_j). \quad (252)$$

We compute $\text{vol}(F_j)$ using the same method, one dimension down. That is, we cone all the facets of F_j to the point w_{j0} . It turns out that the

polyhedra $\{F_{ij}\}$ in the subdivision of F are either tetrahedra or pyramids with quadrilateral base. In case F_{ij} is a tetrahedron we compute $12\text{vol}(F_{ij})$ by taking the appropriate determinant and doubling the answer. In the other case, we compute 6 times the volume of each of the 4 sub-tetrahedron of F_{ij} obtained by omitting a vertex other than w_{0j} and then we add up these volumes. This computes $12\text{vol}(F_{ij})$ regardless of the cyclic ordering of the vertices around the base of F_{ij} .

When we add up all these contributions, we get $12 \text{ vol}(F_j)$. So, our final answer is $48 \text{ vol}(V)$. The reason we scale things up is that we want to have entirely integer quantities.

Potential Overflow Error: With regard to the volume method, a straightforward application of our method can cause an overflow error. We do the calculations using longs (a 64 bit representation of an integer) and the total number we get by adding up 218 smaller numbers is (barely) too large to be reliably represented. However, we observe that all the numbers we compute are divisible by 480. So, before adding each summand to the list, we divide out by 480. This puts us back into the representable range. Recalling that we have scaled up by 60, then computed 48 times the volume, then divided by 480, our final computation is

$$\Omega = 6^4 10^3 \tag{253}$$

times the true volume.

24.2 The Calculations

We work with 3 partitions and then a few extra polytopes.

The Plaid Partition: The plaid partition, described in §7 has 26 clean polytopes modulo the action of the plaid lattice Λ_1 . For each all i, j, k with $|i|, |j|, |k| \geq 3$, we check that each polytope P_k is disjoint from $\lambda_{i,j,k}(P_\ell)$, where $\lambda_{i,j,k}$ is the standard word in the generators of Λ_1 . Given that all the original polytopes intersect the fundamental domain, and given the sizes of the translations in Λ_1 , this check suffices to show that the orbit $\Lambda(\bigcup P_i)$ consists of polytopes with pairwise disjoint interiors and that the union of our polytopes $P_0 \cup \dots \cup P_{25}$ is contained in a fundamental domain for Λ_1 .

At the same time, we compute that

$$\sum_{i=0}^{25} \text{vol}(P_i) = 8. \tag{254}$$

Again, we compute that the scaled volume is 8Ω . This coincides with the volume of the fundamental domain $[-1, 1]^3 \times [0, 1]$ for Λ_1 .

We now know the following three things:

1. The union of the 26 plaid polytopes is contained in a fundamental domain.
2. The union of the 26 plaid polytopes has the same volume as a fundamental domain.
3. The Λ_1 orbit of the 26 plaid polytopes consists of polytopes having pairwise disjoint interiors.

We conclude from this that the Λ_1 orbit of the plaid polytopes P_0, \dots, P_{25} is a partition of \widehat{X} , as desired.

Remark: In view of the fact that we have already proved the Plaid Master Picture Theorem, these checks are really unnecessary, provided that we have correctly interpreted the geometric description of the plaid polytopes and copied down the points correctly. So, these calculations really serve as sanity checks.

The Graph Partitions: We just deal with the (+) graph partition, because the (−) graph partition is isometric to the (+) partition. We make all the same calculations for the (+) graph partition that we made for the plaid partition, and things come out the same way. Once again, the fact that things work out is a consequence of our Master Picture Theorem from [S0], but we made several changes to the polytopes in this paper. We swapped the first and third coordinates, and also translated. So, these calculations serve as sanity checks.

The Reduced Plaid Triple Partition: The reduced plaid triple partition consists of 218 convex rational polytopes, which we scale up by a factor of 60. The scaled polytopes are all clean. Using the above tests, we check

that the polytopes have pairwise disjoint interiors, and are all contained in the fundamental domain $X_{\Pi} = [-1, 1]^3 \times [0, 1]$. Finally, we check that the sum of the volumes is 8. This verifies that we really do have a partition.

There is one more important check we make. Each plaid triple polytope has a 6 letter code which tells how it is obtained as an intersection of the form $A_{-1} \cap A_0 \cap A_1$, where A_k is a polytope in the partition $F^{(k)}(\mathcal{P})$. Here \mathcal{P} is the plaid partition and F is the curve-following dynamics.

We check that each triple plaid Π_k is contained in the three polytopes that are supposed to contain it, and is disjoint from all the others in the plaid partition and its images under the forward and backward curve following map. This verifies that the plaid triple partition really is as we have defined it.

Nine Graph Doubles: We mentioned in §21 that sometimes it happens that $\Psi(\Pi_k)$ is not contained in a single graph polytope of one of the two partitions, but rather a union of two of them. This happens 19 times for the (+) partition and 19 times for the (−) partition. We call these unions of graph polytopes *doubles*.

We check that each graph double (when scaled up by a factor of 60) is clean, that each graph double indeed contains both of its constituent graph polytopes, and that the volume of the graph double is the sum of the volumes of the two constituents. This shows that the union of the two graph polytopes really is a clean convex polytope.

The Rest of the Calculations Each edge crossing problem either involves showing that some linear functional is positive on a polytope, or else that two clean polytopes have pairwise disjoint interiors, or that one clean polytope is contained in another. We simply run the tests and get the outcomes mentioned above. The same goes for the several recalcitrant cases done in connection with the proof of Statement 5 of the Pixellation Theorem given in the last chapter. Finally, checking that there are no errant edges just amounts to listing out the data and checking that there are no forbidden edge assignments. We also survey the assignments visually and see that there are no errant edges.

24.3 The Computer Program

In this last section I'll give a rough account of some of the main features of the program. The main purpose of this account is to show you the kinds of things the program can do. The later entries in this section are rather sketchy. They are designed to let you know what sorts of things the program can do, but they stop short of giving detailed instructions on how to get the program to do it. The program has its own documentation - every feature is explained - and this should help with the details.

24.3.1 Downloading the Program

My computer program can be downloaded from

<http://www.math.brown.edu/~res/Java/PLAID2.tar>

When you download the file, you get a tarred directory. I untar the directory with the Unix command `tar -xvf PLAID2.tar`. (Your system might be different.) Once this is done, you have a new directory called `PlaidModel`. The program resides in this directory and is spread out among many files.

24.3.2 Running and Compiling

The file `Main.java` is the main file. Assuming that the program is compiled already (and you would know by the presence of many `.class` files) compile the program with the command `javac *.java` and then you run the program with the command `java Main`. All this assumes, of course, that your computer can run Java programs. If everything works, a small and colorful window should pop up. This is the control panel. You can launch the other parts of the program from this window.

24.3.3 What to do first

The control panel has a smaller window which lists 10 pop-up windows. If you click on these buttons, additional windows will pop up. The first button you should press is the **Document** button. This will bring up a window which has information about the program. If you read the documentation, you will see how to operate the program.

24.3.4 Presets

The program has 4 preset modes. When you select the **preset** option on the **main** control panel, you bring up an auxiliary control panel called **presets**. This panel has 4 buttons:

- quasi-isomorphism theorem
- plaid master picture theorem
- graph master picture theorem
- plaid-graph correspondence

If you press one of these buttons, various windows will pop up, and they will be automatically set up to best show the advertised features. Moreover, documentation will appear which gives further instructions and explanations.

There is one irritating feature of the program which I should mention. The **preset** features work best when all the auxiliary pop-up windows are closed. If you press a **preset** when some of the pop-up windows are already open, you run the risk of having duplicate windows open, and this causes problems for the program. When you want to use one of the preset buttons, you should close all the auxiliary windows.

24.3.5 Surveying the Data and Proofs

The files starting **Data** contain all the data for the polytope partitions. The file names give some idea of what the files contain. For instance, **Data-GraphPolytopes.java** contains the coordinates of the graph PET polytopes. Some of the files are harder to figure just from the names, but the files themselves have documentation.

The files starting **Proof** contain all the routines for the proof. Again, the names indicate the tests contained in the files. For instance **ProofVolume.java** contains the volume calculations for the partitions.

Part V

The Distribution of Orbits

25 Unbounded Orbits

25.1 Geometric Limits

We start with an irrational $A \in (0, 1)$ and we seek a good offset $V \in \mathbf{R}^3 \times \{0\}$ such that the tiling defined by $\Theta_{A,V}$ has a fat component – i.e. a curve not contained in the tubular neighborhood of any line. We don't care about the directions of the components of the tiling, so we take the domain of Θ_A to be X_1 . We drop the 4th coordinate for convenience.

Our basic idea is to take geometric limits of the big polygons produced in §3.1. Suppose that $\{A_k\} = \{p_k/q_k\}$ is a sequence of even rational parameters converging to some irrational parameter A . Let Γ_k denote the big polygon associated to A_k .

We set $\Theta_k = \Theta_{A_k}$ for ease of notation. The maps Θ_k converge, uniformly on compact sets, to Θ_A . However, the limit map Θ_A is not useful to us because the distance between the closest point on Γ_k and the origin tends to ∞ with k . Also, $\Theta_A(G)$ will intersect the walls of the partition. We need a more robust kind of limit.

Define the *anchor* of Γ_k to be the set $\langle \Gamma_k \rangle \subset G$ consisting of centers of tiles involved in Γ_k which have the form $(1/2, y)$. Let $Z = \{z_k\}$ be a sequence where $z_k \in \langle \Gamma_k \rangle$. Define

$$\Theta_k^Z(c) = \Theta_k(c + z_k). \quad (255)$$

The map Θ_k^Z behaves near the origin like Θ_k behaves near z_k . Since Θ_k is locally affine, there is some vector V_k such that $\Theta_k^Z = \Theta_{k,V_k}$. Passing to a subsequence, we get some offset V_Z such that

$$\lim_{k \rightarrow \infty} \Theta_k^Z = \Theta_{A,V_Z}. \quad (256)$$

By construction, the paths $\Gamma_k - z_k$ will converge to an unbounded path that involves the tile centered at $(1/2, 1/2)$.

The trick is finding a sequence Z so that V_Z is a good offset. One such sequence is given by $z_k = (1/2, \tau_k + 1/2)$, where τ_k is as in §1.1. However, this choice never leads to a good offset. Our idea is to pick a rational approximation so that the sets $\langle \Gamma_k \rangle$ are huge in the sense that the finite set $\Theta_k(\langle \Gamma_k \rangle)$ becomes dense in a Cantor set as $k \rightarrow \infty$. This structure will allow us to get good offsets in the limit, and the limiting tiling will turn out to have a fat path.

25.2 Approximating Irrationals

Our first step is to find a good sequence $\{p_n/q_n\}$ of even rationals that limits to A . The sequence we choose is related to, but distinct from, the sequence of continued fraction approximants. (The continued fraction approximants are not always even rationals.) We defer the proofs of the technical lemmas to the next chapter.

Even Predecessors: Two rationals a_1/b_1 and a_2/b_2 are called *Farey related* if $|a_1b_2 - a_2b_1| = 1$. Let p'/q' and p/q be two even rational parameters. We write $p'/q' \leftarrow p/q$ if p/q and p'/q' are Farey related and $\omega' < \omega$. Here $\omega = p + q$ as usual. We call p'/q' the *even predecessor* of p/q . This rational is unique.

Core Predecessors: Let $\tau \in (0, \omega/2)$ be as in §1.1. We define $\kappa \geq 0$ to be the integer so that

$$\frac{\kappa}{2\kappa + 1} \leq \frac{\tau}{\omega} < \frac{\kappa + 1}{2(\kappa + 1) + 1}. \quad (257)$$

We only get equality on the left hand side when $p/q = 1/2n$. In this case, $\tau = 1$. Define

$$\hat{p} = p - 2\kappa p', \quad \hat{q} = q - 2\kappa q'. \quad (258)$$

We will see in the next chapter that \hat{p}/\hat{q} is an even rational in $(0, 1)$. We call \hat{p}/\hat{q} the *core predecessor* of p/q . Note that when $\kappa = 0$ we have $\hat{p}/\hat{q} = p/q$. We only care about the core predecessor when $\kappa \geq 1$.

Predecessors: Given an even rational parameter p/q , we define the *predecessor* p^*/q^* as follows:

- If $p = 1$ then $p^*/q^* = 0/1$.
- If $p \geq 2$ and $\kappa = 0$ then $p^*/q^* = p'/q'$, the even predecessor of p/q .
- If $p \geq 2$ and $\kappa \geq 1$ then $p^*/q^* = \hat{p}/\hat{q}$, the core predecessor of p/q .

We write $p^*/q^* \prec p/q$. This definition turns out to be very well adapted to the plaid model. Lemma 26.1 from the next chapter collects together many of the relations between these rationals.

The Predecessor Sequence: In the next chapter, we will prove the following result.

Lemma 25.1 *Let $A \in (0, 1)$ be irrational. Then there exists a sequence $\{p_k/q_k\}$ such that*

- $p_0/q_0 = 0/1$
- $p_k/q_k \prec p_{k+1}/q_{k+1}$ for all k
- $A = \lim p_k/q_k$.

We call $\{p_k/q_k\}$ the *predecessor sequence*. This terminology suggests that the predecessor sequence is unique. However, we will not bother to prove this. We just need existence, not uniqueness.

Diophantine Result: Let $\{p_k/q_k\}$ be the predecessor sequence converging to A . We classify a term p_k/q_k in the predecessor sequence as follows:

- weak: $\tau_{k+1} < \omega_{k+1}/4$. Here $\kappa_{k+1} = 0$.
- strong: $\tau_{k+1} \in (\omega_{k+1}/4, \omega_{k+1}/3)$. Here $\kappa_{k+1} = 0$.
- core: $\tau_{k+1} > \omega_{k+1}/3$. Here $\kappa_{k+1} \geq 1$.

Lemma 25.2 *The predecessor sequence has infinitely many non-weak terms. For each non-weak term p_k/q_k , we have*

$$\left| A - \frac{p_k}{q_k} \right| < \frac{48}{q_k^2}.$$

The Approximating Sequence: We define the approximating sequence to be the set of terms p_k/q_k in the predecessor sequence such that either

- p_k/q_k is core.
- p_k/q_k is strong and p_{k-1}/q_{k-1} is not core.

If there are infinitely many core terms in the predecessor sequence, then the approximating sequence contains all of these. If there are only finitely many core terms in the predecessor sequence, then there are infinitely many strong terms, and the approximating sequence contains all but finitely many of these. So, in all cases, the approximating sequence is an infinite sequence.

25.3 Arc Copying

Let $R_{p/q}$ denote the rectangle bounded by the bottom, left, and top of the first block, and by whichever vertical line of capacity at most 4 is closest to the left edge of the first block. Let $\gamma_{p/q}$ denote the subset of $\Gamma_{p/q}$ contained in the box $R_{p/q}$. When the dependence on the parameter is implied we will suppress it from our notation. In subsequent chapters, we prove the following result.

Lemma 25.3 (Box) *For any even rational parameter p/q , the set $\gamma_{p/q}$ is an arc whose endpoints lie on the right edge of $R_{p/q}$.*

Figure 8.2 shows the big polygons associated to two different parameters. Notice that the polygon $\Gamma_{12/29}$ copies some of $\Gamma_{5/12}$. The boxes $B_{5/12}$ and $B_{12/29}$ are the first columns (i.e. the union of the leftmost 3 sub-rectangles) of the tic-tac-toe grids shown in each of the two pictures.

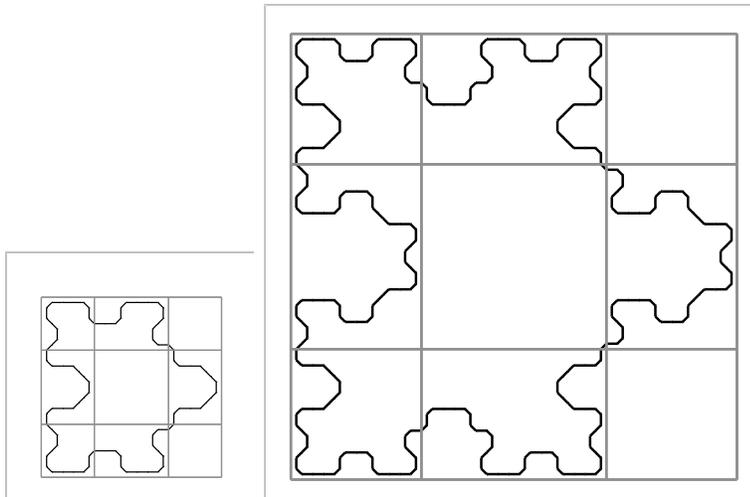


Figure 8.2: Arc copying for 5/12 and 12/29.

Let TH and BH denote the top and bottom horizontal lines of capacity 2 with respect to some rational parameter. In the statement of the results below, it will be clear which parameters these lines depend on. In the subsequent chapters we prove the following result.

Theorem 25.4 (Copy) *Let $p_0/q_0, p_1/q_1$ be two successive terms in the approximating sequence. Then there is some vertical translation Υ such that $\Upsilon(R_0)$ is contained below the horizontal midline of R_1 , and $\Upsilon(\gamma_0) \subset \gamma_1$. Moreover, either $\Upsilon(BH_0) = BH_1$ or $\Upsilon(TH_0) = BH_1$.*

The Copy Theorem has a more symmetric formulation. Let $\eta = \omega - 2\tau$. Geometrically η is the vertical distance between lines BH and TH .

Corollary 25.5 *Let $p_0/q_0, p_1/q_1$ be two successive terms in the approximating sequence. Then there are vertical translations Υ_B and Υ_T such that*

$$\Upsilon_B(\gamma_0) \subset \gamma_1, \quad \Upsilon_T(\gamma_0) \subset \gamma_1.$$

The image $\Upsilon_B(\gamma_0)$ lies below the horizontal midline of B_1 and The image $\Upsilon_T(\gamma_0)$ lies above the horizontal midline of B_1 . Furthermore

$$\Upsilon_T(*) - \Upsilon_B(*) = (0, \eta_1 \pm \eta_0).$$

Proof: We set $\Upsilon_B = \Upsilon$ from the Copy Theorem. Let R_j denote reflection in the horizontal midline of B_j . Note that $R_j(\gamma_j) = \gamma_j$. We let

$$\Upsilon_T = R_1 \circ \Upsilon_B \circ R_2.$$

By symmetry $\Upsilon_T(\gamma_0) \subset \gamma_1$ and the image lies above the horizontal midline of B_1 .

If $\Upsilon_B(BH_0) = BH_1$ then $\Upsilon_T(TH_0) = TH_1$. In this case

$$\Upsilon_T(*) - \Upsilon_B(*) = \eta_1 - \eta_0.$$

If $\Upsilon_B(TH_0) = BH_1$ then $\Upsilon_T(BH_0) = TH_1$. In this case

$$\Upsilon_T(*) - \Upsilon_B(*) = \eta_1 + \eta_0.$$

Either case gives us the final statement of the corollary. ♠

Corollary 25.6 *$\langle \Gamma_1 \rangle$ contains two copies of $\langle \Gamma_0 \rangle$, separated by one of the two vectors $(0, \eta_1 \pm \eta_0)$.*

We can iterate this result. Let Γ_k be the big polygon corresponding to the k th term in the approximation sequence.

Corollary 25.7 *$\langle \Gamma_k \rangle$ contains 2^k points of the form*

$$c_k + \sum_{i=0}^{k-1} \epsilon_i(0, d_i),$$

where d_i is one of the two terms $\eta_{i+1} - \eta_i$ for each i . Here c_k is some integer and $\{\epsilon_i\}$ is any binary sequence of length k .

25.4 The Images of the Anchors

In this section we prove a result about the images of the anchors. We keep the notation from the previous section. First we make some preliminary observations. For any parameter A , a fundamental domain for X_1 is $[-1, 1]^3$. The coordinates on this space are (T, U_1, U_2) . We think of this space as a fibration over the T coordinate. The (U_1, U_2) fibers are all flat square tori. Let γ_P denote the geodesic in the fiber over $T = P$ which has slope -1 . Here $P = 2A/(1 + A)$ as usual. To simplify notation, we set $\gamma_k = \gamma_{P_k}$.

Let $Y = \{1/2\} \times \mathbf{Z}$. Using the formula for Θ_k we see that $\Theta_k(Y) \subset \gamma_k$. Define

$$I_k = \Theta_k(\langle \Gamma_k \rangle) \subset \gamma_k \quad (259)$$

The last containment comes from the fact that $\langle \Gamma_k \rangle \subset Y$, by definition. From Corollary 25.7 we know that I_k has at least 2^k points, namely

$$\Theta_k(c_k) + \sum_{i=0}^{k-1} \epsilon_i \Theta_k(0, d_i), \quad \epsilon_0, \dots, \epsilon_{k-1} \in \{0, 1\}. \quad (260)$$

The addition on the right takes place in the flat square torus above the fiber $T = P_k$.

Now we are going to analyze the terms in I_k and show that, on a subsequence, the sequence of sets $\{I_k\}$ converges to a Cantor set. We will interpret this Cantor set as a subset of the available offsets we can get by taking the kind of limit discussed in §25.1.

Lemma 25.8 *As $k \rightarrow \infty$, the vectors $\Theta_k(0, d_i)$ converge to $\Theta_A(0, d_i)$.*

Proof: We have

$$\Theta_k(0, d_i) = (2d_i, 0, 2d_i P_k) \equiv (0, -d_i P_k, +d_i P_k) \pmod{\Lambda_1}. \quad (261)$$

Using Lemma 25.2 and the fact that the map $x \rightarrow 2x/(x + 1)$ is 2-lipschitz on $(0, 1)$, we have

$$\|P_k - P\| \leq \frac{96}{\omega_k^2}. \quad (262)$$

The lemma follows immediately from this estimate and from the fact that $|d_i| \leq 2\omega_k$ once $k \geq i$. ♠

The vectors $\Theta_A(d_i)$ are growing unboundedly long, but mod Λ_1 these vectors converge to 0 exponentially fast. That is the same kind of convergence we are after here.

Lemma 25.9 $\|\Theta_A(0, d_i)\| < C2^{-i}$ mod Λ_1 for some constant C .

Proof: By the Copy Theorem, $\omega_k < \omega_{k+1}/2$. Hence $\omega_i \geq 2^i$. So, it suffices to prove that

$$\|\Theta_A(0, d_i)\| < \frac{C}{\omega_i},$$

for some constant C . By the triangle inequality and Equation 262, it suffices to prove that

$$\|\Theta_i(0, \eta_i)\| < \frac{C}{\omega_i}.$$

We have

$$\Theta_i(0, \eta_i) = (0, -P_i\eta_i, P_i\eta_i). \quad (263)$$

We have

$$\eta_i P_i = (\omega_i - 2\tau_i)P_i = 2p_i - \frac{4\tau_i p_i}{\omega_i} = 2p_i - 2a_i - \frac{2}{\omega_i}.$$

Here a_i is some integer. This means that

$$P_i\eta_i = \frac{2}{\omega_i} \pmod{2\mathbf{Z}}. \quad (264)$$

Our result follows from Equations 263 and 264. ♠

Combining the results above, we can say that I_k is becoming dense in the Cantor set

$$\sum_{i=0}^{\infty} \epsilon_i \Theta_A(0, d_i), \quad \epsilon_i \in \{0, 1\}. \quad (265)$$

Every infinite sum makes sense because the terms decay exponentially fast. The sequence of translations is given by $\{\Theta_A(c_k)\}$.

In the next section we use this result to finish the proof of the Unbounded Orbits Theorem.

25.5 The End of the Proof

Here we give a criterion for V to be a good offset. Define

$$\mathbf{Q}[P] = \{r_1 + r_2P \mid r_1, r_2 \in \mathbf{Q}\}. \quad (266)$$

Lemma 25.10 *If $V = (P, u_1, u_2)$ with $u_1, u_2 \notin \mathbf{Q}[P]$, Then V is a good offset.*

Proof: The (U_1, U_2) fiber over and point of $\mathbf{Q}[P]$ intersects the walls of our partition in rectangles bounded by lines defined over $\mathbf{Q}[P]$. This is a consequence of the fact that the 4-dimensional polytopes of our partition are integral. Given $c \in G$, which has half-integer coordinates $(m + 1/2, n + 1/2)$, we compute

$$\Theta_{A,V}(c) = (2mP + P, 2mP + P, 2mP + P) + (2n + 1, 0, 2nP + P) + V. \quad (267)$$

The first coordinate always lies in $\mathbf{Q}[P]$ and the second two coordinates never do. Hence, this point never lies in the boundary of the partition. ♠

Now we go back to our idea of choosing $z_k \in \langle \Gamma_k \rangle$ and taking a geometric limit of the maps Θ_k^Z where $Z = \{z_k\}$.

Lemma 25.11 *There exists a sequence $Z = \{z_k\}$ with $z_k \in \langle \Gamma_k \rangle$ such that the some subsequential limit V_Z is a good offset.*

Proof: Passing to a subsequence, we can assume that $\Theta_A(c_k)$ converges mod Λ_1 to come vector V . On this subsequence, the sets I_k are becoming dense in the Cantor set

$$\Omega = V + \sum_{i=0}^{\infty} \epsilon_i \Theta_A(0, d_i), \quad \epsilon_i \in \{0, 1\}. \quad (268)$$

Ω lies the geodesic γ_p , which lies in the fiber over $T = P$ and has slope -1 in the (U_1, U_2) plane. In particular, the set in Equation 268 contains points of the form (P, u_1, u_2) , where neither u_1 nor u_2 belongs to $\mathbf{Z}[P]$. In other words, according to Lemma 25.10, the Cantor set Ω contains vectors which are good offsets. But this means that there exists a sequence Z such that $\Theta_k(z_k) \rightarrow V_Z$ where V_Z is a good offset. But $\Theta_k(z_k) = \Theta_k^Z(0, 0)$. In other

words, V_Z is a good offset. ♠

We choose such a sequence and take a limit. Let $\Theta_{A,V}$ be the limiting map. Since V is a good offset, the tiling associated to $\Theta_{A,V}$ is entirely well defined. By construction, the curve Γ_∞ of the tiling through $(1/2, 1/2)$ is unbounded. However, we can say more than this. There are infinitely many indices for which some initial portion of Γ_∞ agrees with $\Gamma_k \cap R_k$. Here R_k is the box from the Copy Lemma. But this means that the projection of Γ_∞ onto the x -axis exceeds $p_k/2$ for infinitely many k . Hence the projection of Γ_∞ onto the x -axis is unbounded. At the same time, $\langle \Gamma_\infty \rangle$ is infinite. This is to say that Γ_∞ contains infinitely many points which are within $1/2$ of the y -axis. These two properties together imply that Γ_∞ is fat. This completes the proof of the Unbounded Orbits Theorem.

26 Some Elementary Number Theory

In this chapter we will justify all the statements quoted in §25.2. We will also prove a number of other technical results which establish useful identities and inequalities between the rationals discussed in §25.2.

26.1 All About Predecessors

We fix an even rational parameter p/q . To avoid trivialities, we assume that $p > 1$. We recall the following definitions from §25.2.

- $\omega = p + q$.
- $\tau \in (0, \omega/2)$ is the unique solution to $2p\tau \equiv \pm 1 \pmod{\omega}$.
- $\kappa \in \{0, 1, 2, \dots\}$ is such that

$$\frac{\kappa}{2\kappa + 1} \leq \frac{\tau}{\omega} < \frac{\kappa + 1}{2(\kappa + 1) + 1}. \quad (269)$$

- p'/q' is the even predecessor of p/q . That is, $|pq' - qp'| = 1$ and $q < q'$.
- \hat{p}/\hat{q} is the core predecessor of p/q . Here $\hat{p} = p - 2\kappa p'$ and $\hat{q} = p - 2\kappa q'$.

We mean to define ω, τ, κ for the other parameters as well. Thus, for instance, $\omega' = p' + q'$.

Lemma 26.1 *The following is true.*

1. $\hat{p}/\hat{q} \in (0, 1)$ is an even rational in lowest terms.
2. Either $\tau - \tau' = \kappa\omega'$ or $\tau + \tau' = (1 + \kappa)\omega'$. In all cases, $\tau' \leq \tau$.
3. p'/q' is the even predecessor of \hat{p}/\hat{q} .
4. $\hat{\kappa} = 0$.
5. $\omega - 2\tau = \hat{\omega} - 2\hat{\tau}$.
6. If $\kappa = 0$ then $\tau' = \tau$ when $\tau < \omega/4$ and $\tau' = \omega' - \tau$ when $\tau > \omega/4$.
7. $\hat{\omega} \leq (3/(3 + 2\kappa))\omega$.

26.1.1 Statement 1

We first give a formula for p'/q' . There is some integer $\theta > 0$ so that

$$2p\tau = \theta(p + q) \pm 1. \quad (270)$$

Since $q > 1$ we have $\theta < 2\tau$. We also have $\theta \leq p$. The case $\theta = p$ forces $p = 1$. Hence $\theta < \min(p, 2\tau)$. Rearranging Equation 270 we get

$$|pq'' - qp''| = 1, \quad p'' = \theta, \quad q'' = 2\tau - \theta.$$

This implies that p''/q'' is in lowest terms. Moreover, both p'' and q'' are odd. Hence

$$p' = p - p'' = p - \theta, \quad q' = q - q'' = q - 2\tau + \theta. \quad (271)$$

This gives us

$$\omega' = \omega - 2\tau \leq \omega - 2\omega \frac{\kappa}{2\kappa + 1} = \frac{\omega}{2\kappa + 1}. \quad (272)$$

Rearranging this, we get

$$2\kappa p' + 2\kappa q' + (p' + q') \leq p + q. \quad (273)$$

Given that $|pq' - qp'| = 1$, Equation 273 forces $2\kappa p' < p$ and $2\kappa q' < q$. Hence $\hat{p} = p - 2\kappa p' > 0$ and $\hat{q} = q - 2\kappa q' > 0$. Since \hat{p}/\hat{q} is Farey related to $p'/q' \in (0, 1)$, we see that \hat{p}/\hat{q} also lies in $(0, 1)$ and is in lowest terms. Since p/q is an even rational, so is \hat{p}/\hat{q} . This proves Statement 1.

26.1.2 Statement 2

We use the formulas derived in the previous section. Let θ be as above. Choosing the sign as in Equation 270, we have

$$\theta\omega' \pm 1 = \theta(\omega - 2\tau) \pm 1 = (\theta\omega \pm 1) - 2\theta\tau = 2p\tau - 2\theta\tau = 2p'\tau. \quad (274)$$

This shows that $2p'\tau \equiv \pm 1 \pmod{\omega'}$. But, by definition, $2p'\tau' \equiv \pm 1 \pmod{\omega'}$. Hence $2p'\tau \equiv \pm 2p'\tau' \pmod{\omega'}$. Since $2p'$ is relatively prime to ω' we see that

$$\tau \equiv \pm\tau' \pmod{\omega'}. \quad (275)$$

If $\tau \equiv \tau' \pmod{\omega'}$ then $\tau' = \tau - \kappa\omega'$ provided that this expression lies in $(0, \omega')$. We compute

$$\tau - \kappa\omega' = \tau - \kappa(\omega - 2\tau) = (2\kappa + 1)\tau - \kappa\omega > 0 \quad (276)$$

$$\tau - (\kappa + 1)\omega' = \tau - (\kappa + 1)(\omega - 2\tau) = (2\kappa + 3)\tau - (\kappa + 1)\omega < 0 \quad (277)$$

These two inequalities show that $\tau - \kappa\omega' \in (0, \omega')$. In this case it is obvious that $\tau' \leq \tau$.

When $\tau \equiv -\tau \pmod{\omega}$, a similar argument shows that

$$(\kappa + 1)\omega' - \tau \in (0, \omega'),$$

and this forces $\tau' = (\kappa + 1)\omega' - \tau$. In this case, the two inequalities $\tau + \tau' \geq \omega'$ and $\tau' < \omega'/2$ establish $\tau' < \tau$.

26.1.3 Statements 3 and 5

Let \widehat{p}/\widehat{q} be the core predecessor of p/q . Since $\widehat{p} = p - 2\kappa p'$ and $\widehat{q} = q - 2\kappa q'$ the two rationals p'/q' and \widehat{p}/\widehat{q} are both Farey related. Equation 272, which is a strict inequality when $p > 1$, says that $(2\kappa + 1)\omega' < \omega$. Since $\widehat{\omega} = \omega - 2\kappa\omega'$, we conclude that $\omega' < \widehat{\omega}$. Hence p'/q' is the even predecessor of \widehat{p}/\widehat{q} .

Since p'/q' is the even predecessor of both p/q and \widehat{p}/\widehat{q} , Equation 272 tells us that $\omega - 2\tau = \omega' = \widehat{\omega} - 2\widehat{\tau}$. Hence $\omega - 2\tau = \widehat{\omega} - 2\widehat{\tau}$.

26.1.4 Statement 4

If Statement 4 is false, then Equation 272, applied to the pair $(\widehat{p}/\widehat{q}, p'/q')$, gives

$$\widehat{\omega} - 3\omega' > 0. \quad (278)$$

Combining this with the fact that

$$\widehat{\omega} = \omega - 2\kappa\omega' > 0,$$

we can say that

$$\omega - (2\kappa + 3)\omega' > 0.$$

But then we have

$$\frac{1}{2\kappa + 3} < 1 - 2\left(\frac{\tau}{\omega}\right) = \frac{\omega'}{\omega} < \frac{1}{2\kappa + 3}.$$

This is a contradiction.

26.1.5 Statement 6

When $\kappa = 0$, Statement 1 says that either $\tau = \tau'$ or $\tau + \tau' = \omega'$. Also, from Equation 272 we see that $2\omega' < \omega$ if and only if $\tau > \omega/4$.

Suppose that $\tau' = \tau$ and $\tau > \omega/4$. Then

$$\tau = \tau' < \omega'/2 < \omega/4.$$

This is a contradiction. So, $\tau' = \tau$ implies that $\tau < \omega/4$. Suppose that $\tau' = \omega' - \tau$ and $\tau < \omega/4$. Then

$$2\tau > 2\tau' = 2\omega' - 2\tau > \omega - 2\tau.$$

This gives $\tau > \omega/4$, a contradiction. So, $\tau' = \omega - \tau$ implies that $\tau > \omega/4$.

26.1.6 Statement 7

We have

$$\widehat{\omega} = \omega - 2\kappa\omega' = \omega - 2\kappa(\omega - 2\tau) = (1 - 2\kappa)\omega + 4\kappa\tau.$$

By definition

$$\frac{\tau}{\omega} < \frac{\kappa + 1}{2\kappa + 3}.$$

Hence

$$\frac{\widehat{\omega}}{\omega} = 1 - 2\kappa + 4\kappa\frac{\tau}{\omega} \leq 1 - 2\kappa + 4\kappa\frac{\kappa + 1}{2\kappa + 3} = \frac{3}{3 + 2\kappa}.$$

26.2 Existence of the Predecessor Sequence

We begin by proving the existence of an auxiliary sequence. The notation $p'/q' \leftarrow p/q$ means that p'/q' is the even predecessor of p/q .

Lemma 26.2 *Let $A \in (0, 1)$ be irrational. There exists a sequence $\{p_n/q_n\}$ converging to A such that $p_n/q_n \leftarrow p_{n+1}/q_{n+1}$ for all n .*

Proof: Let \mathbf{H}^2 denote the upper half-plane model of the hyperbolic plane. We have the usual Farey triangulation of \mathbf{H}^2 by ideal triangles. The geodesics bounding these triangles join rationals p_1/q_1 and p_2/q_2 such that $|p_1q_2 - p_2q_1| = 1$. We call these geodesics the *Farey Geodesics*. Two of the Farey geodesics in the tiling join the points $0/1$ and $1/1$ to $1/0$. This last point

is interpreted as the point at infinity in the upper half plane model of \mathbf{H}^2 . Let $Y = \{(x, y) \mid 0 < x < 1\}$ be the open portion of \mathbf{H}^2 between these two vertical geodesics.

Each Farey geodesic in Y which joins two even rationals p/q and p'/q' can be oriented so that it points from p/q to p'/q' if and only if $p'/q' \leftarrow p/q$. We leave the rest of the Farey geodesics unoriented.

We claim that there is a backwards oriented path from $0/1$ to A . To see this, choose any sequence of even rationals converging to A and consider their sequence of even predecessors. This gives us a sequence of finite directed paths joining $0/1$ to rationals which converge to A . Given the local finiteness of the Farey triangulation – meaning that any compact subset of \mathbf{H}^2 intersects only finitely many triangles – we can take a limit of these paths, at least on a subsequence, and get a directed path converging to A .

Reading off the vertices of this path gives us a sequence $\{p_n/q_n\}$ with $p_n/q_n \leftarrow p_{n+1}/q_{n+1}$ for all n . ♠

Proof of Lemma 25.1: We get the predecessor sequence from the sequence in Lemma 26.2 in the following way. Before each term p_n/q_n with $\kappa_n \geq 1$ we insert the rational $\widehat{p}_n/\widehat{q}_n$. This is the core predecessor of p_n/q_n . We know from the lemma above that $\widehat{\kappa}_n = 0$ and that p_{n-1}/q_{n-1} is the predecessor of $\widehat{p}_n/\widehat{q}_n$. In short

$$p_{n-1}/q_{n-1} \prec \widehat{p}_n/\widehat{q}_n \prec p_n/q_n.$$

Once we make all these insertions, the resulting sequence is a predecessor sequence which converges to A . Again, we will not stop to prove uniqueness because we don't care about it. ♠

26.3 Existence of the Approximating Sequence

We call the sequence from Lemma 26.2 the *even predecessor sequence*. As we have just remarked, the even predecessor sequence is a subsequence of the predecessor sequence.

Lemma 26.3 *Let $\{p_n/q_n\}$ be the even predecessor sequence which converges to A . There are infinitely many values of n such that $\tau_{n+1} > \omega_{n+1}/4$.*

Proof: If this lemma is false, then we might chop off the beginning and assume that this never happens. We have the formula $\omega_n = \omega_{n+1} - 2\tau_{n+1}$. So,

if this lemma is false then we have $2\omega_n < \omega_{n+1}$ only finitely often. Chopping off the beginning of the sequence, we can assume that this never occurs.

We never have $2\omega_n = \omega_{n+1}$ because ω_{n+1} is odd. So, we always have $\omega_n > \omega_{n+1}$. In this case, we have

$$r_{n-1} = 2r_n - r_{n+1}$$

for $r \in \{p, q\}$. Applying this iteratively, we get that $r_k = (k+1)r_1 - kr_0$ for $k = 1, 2, 3, \dots$. But then $\lim p_n/q_n = (p_1 - p_0)/(q_1 - q_0)$, which is rational. This is a contradiction.

Now let p_n/q_n be a term with $2\omega_n < \omega_{n+1}$. We will consider the case when $p_n/q_n < p_{n+1}/q_{n+1}$. The other case has a similar treatment. We introduce the new rational

$$p_n^*/q_n^* = \frac{p_{n+1} - p_n}{q_{n+1} - q_n}.$$

Note that the three rationals p_n/q_n and p_{n+1}/q_{n+1} and p_n^*/q_n^* form the vertices of a Farey triangle. Moreover $p_{n+1}/q_{n+1} < p_n^*/q_n^*$ because p_{n+1}/q_{n+1} is the Farey sum of p_n/q_n and p_n^*/q_n^* .

Given the geometry of the Farey graph, we have

$$A \in [p_n/q_n, p_n^*/q_n^*].$$

Moreover

$$|p_n/q_n - p_n^*/q_n^*| = \frac{1}{q_n q_n^*} \leq \frac{1}{q_n^2} < \frac{4}{\omega_n^2}.$$

The second-to-last inequality comes from the fact that $q_n^* > q_n$. ♠

Proof of Lemma 25.2: Let $\{p_k/q_k\}$ be the even predecessor sequence. If it happens infinitely often that $\kappa_k \geq 1$ then there are infinitely many core terms in the predecessor sequence. Otherwise the predecessor sequence and the even predecessor sequence have the same tail end. But then the previous result shows that there are infinitely many values of k for which $2\omega_k < \omega_{k+1}$. This gives infinitely many strong terms. Suppose that $p/q, p^*/q^*$ are two consecutive terms in the predecessor sequence, with p/q non-weak.

Case 1: Suppose that p/q and p^*/q^* are both terms in the even sequence. Then $p/q \leftarrow p^*/q^*$ and $\tau_* > \omega_*/4$. But then $2\omega < \omega_*$, as in the proof of Lemma 26.3. From here, we get the Diophantine estimate from Lemma 26.3.

Case 2: Suppose that neither p/q nor p^*/q^* are terms in the even sequence. This case cannot happen because when p^*/q^* is the core predecessor of some rational and p/q is the core predecessor of p^*/q^* . This contradicts Statement 4 of Lemma 26.1.

Case 3: Suppose that p/q is a term of the even sequence but p^*/q^* is not. Then the term in the predecessor sequence after p^*/q^* is p^{**}/q^{**} , which belongs to the even sequence. By Statement 3 of Lemma 26.1, we have $p/q \leftarrow p^{**}/q^{**}$. Moreover $2\omega < \omega^* < \omega^{**}$. From here, we get the Diophantine estimate from Lemma 26.3.

Case 4: Suppose that p/q is not a term in the even sequence but p^*/q^* is. Then p/q is the core predecessor of p^*/q^* . Statement 4 of Lemma 26.1 says that $\kappa = 0$. Hence the term p'/q' preceding p/q in the predecessor sequence is the even predecessor of p/q . Since $\kappa = 0$, we have $\tau < \omega/3$. Using the formula $\omega' = \omega - 2\tau$, we see that $3\omega' > \omega$. At the same time $\omega' < \omega < 2\omega^*$. Lemma 26.3 gives us

$$\left| A - \frac{p'}{q'} \right| < \frac{4}{(\omega')^2} < \frac{36}{\omega^2}.$$

At the same time

$$\left| \frac{p}{q} - \frac{p'}{q'} \right| = \frac{1}{qq'} < \frac{4}{\omega\omega'} < \frac{12}{\omega^2}.$$

The result in this case then follows from the triangle inequality. ♠

26.4 Another Identity

Let p/q be an even rational parameter with $\kappa \geq 1$. Let \widehat{p}/\widehat{q} be the core predecessor. Define

$$w = \widehat{\tau}, \quad h = \omega - 2\tau. \tag{279}$$

Here is an identity between these quantities.

Lemma 26.4

$$(2\kappa + 1)h + 2w = \omega. \tag{280}$$

Proof: Define

$$F_\kappa = \frac{\kappa}{2\kappa + 1}. \quad (281)$$

Let F be the union of all such rationals.

We rescale the first block so that it coincides with the unit square. This rescaling amounts to dividing by ω . For each quantity λ defined above, we let $\lambda^* = \lambda/\omega$. Within the unit square, our identity is

$$(2\kappa + 1)h^* + 2w^* = 1. \quad (282)$$

The rescaled horizontal and vertical lines depend continuously on the parameter τ^* . Moreover, when $\tau^* \in (F_\kappa, F_{\kappa+1})$ is rational, all the horizontal and vertical lines of capacity up to $4\kappa + 2$ are distinct, because the denominator of the corresponding rational is at least $2\kappa + 1$. Therefore, κ is a locally constant function of τ^* , and changes only when τ^* passes through a value of F .

Consider what happens when $\tau^* = F_\kappa + \epsilon$, and $\epsilon \in (0, F_{\kappa+1} - F_\kappa)$. We compute

$$(2\kappa + 1)\tau^* - \kappa = (2\kappa + 1)\epsilon < 1/2 \quad (283)$$

Hence $w = (2\kappa + 1)\epsilon$, which means that $2w^* = (4\kappa + 2)\epsilon$. We also have $h^* = 1 - 2\tau^*$. From these equations one can easily compute that Equation 282 holds for all $\epsilon \in (0, F_{\kappa+1} - F_\kappa)$. ♠

Lemma 26.5 $w = \widehat{\tau}$. In other words, the line $x = \widehat{\tau}$ has capacity $4\kappa + 2$ with respect to p/q .

Proof: In view of Equation 280, we just have to prove that

$$(2\kappa + 1)(\omega - 2\tau) + 2\widehat{\tau} = \omega \quad (284)$$

By Statement 5 of Lemma 26.1 we have $\omega - 2\tau = \widehat{\omega} - 2\widehat{\tau}$. But then Equation 284 is equivalent to

$$(2\kappa + 1)\widehat{\omega} - 4\kappa\tau = \omega. \quad (285)$$

This equation holds, because

$$\omega - \widehat{\omega} = 2\kappa\omega' = 2\kappa(\omega - 2\tau) = 2\kappa(\widehat{\omega} - 2\widehat{\tau}) = 2\kappa\widehat{\omega} - 4\kappa\widehat{\tau}.$$

This completes the proof. ♠

27 The Box Lemma and The Copy Theorem

27.1 The Weak and Strong Copying Lemmas

In this chapter we reduce the Box and Copy Lemmas from §25.3 to three other results, the Weak Copy Lemma, the Strong Copy Lemma, and the Core Copy Lemma. We explain the Weak and Strong Copy Lemmas in this section and the Core Copy Lemma in the next.

Let p/q be an even rational parameter with $\kappa = 0$. Let p'/q' be the even predecessor of p/q . Let Π and Π' denote the plaid tilings with respect to these two parameters. With $\kappa = 0$ there are two subcases.

- **weak:** $2\omega' > \omega$. This corresponds to $\tau < \omega/4$.
- **strong:** $2\omega' < \omega$. This corresponds to $\tau > \omega/4$ (and $\tau < \omega/3$.)

Let $R_{p'/q'}$ be the rectangle associated to the parameter p'/q' in §25.3. Again, this rectangle is bounded by the lines

- $y = 0$.
- $y = \omega'$.
- $x = 0$.
- $x = \min(\tau', \omega - 2\tau')$.

The right side of $R'_{p'/q'}$ is whichever line of capacity at most 4 is closest to the y -axis.

In the weak case, we let Σ' denote the subset of $R_{p'/q'}$ bounded above by the line

$$y = \omega' - \min(\tau', \omega' - 2\tau'). \quad (286)$$

That is, the top of Σ' is whichever horizontal line of capacity at most 4 lies closest to the top of $R_{p'/q'}$.

In the strong case, we let $\Sigma' = R_{p'/q'}$.

In both cases we define

$$\Sigma = \Sigma'. \quad (287)$$

Even though $\Sigma = \Sigma'$ it is useful to have separate notation, so that in general an object X' corresponds to the parameter p'/q' and an object X corresponds to the parameter p/q .

In the next chapter we will prove the following result.

Lemma 27.1 (Weak and Strong Copying) $\Sigma' \cap \Pi' = \Sigma \cap \Pi$.

27.2 The Core Copy Lemma

Let p/q be an even rational parameter with $\kappa \geq 1$. Let \widehat{p}/\widehat{q} be the core predecessor of p/q . Let Π and $\widehat{\Pi}$ denote the plaid tilings with respect to these two parameters.

Let Υ be vertical translation by $(\omega + \widehat{\omega})/2$. It follows from Statement 5 of Lemma 26.1 that Υ maps the horizontal lines of capacity ± 1 w.r.t \widehat{p}/\widehat{q} to the lines of capacity ± 1 w.r.t. p/q .

Let $R_{\widehat{p}/\widehat{q}}$ is the rectangle associated to the parameter \widehat{p}/\widehat{q} in §25.3. Define

$$\widehat{\Sigma} = R_{\widehat{p}/\widehat{q}}, \quad \Sigma = \Upsilon(\widehat{\Sigma}). \quad (288)$$

In §9 we will prove the following result.

Lemma 27.2 (Core Copying) $\Sigma \cap \Pi = \Upsilon(\widehat{\Sigma} \cap \widehat{\Pi})$.

27.3 Proof of the Box Lemma

Our proof goes by induction on the denominator of the parameter. We will suppose that p/q is a rational with the smallest denominator for which we don't know the truth of the result.

By construction, Γ can only intersect ∂R in the right edge, which we call ρ . If we knew that Γ intersects ρ twice, then, because Γ is a closed loop, we could conclude that the portion of Γ contained in Σ is an arc whose endpoints are on ρ . If ρ has capacity 2, then Γ can intersect ρ at most twice, and the large x -diameter of Γ implies that Γ does intersect ρ . By symmetry, Γ intersects ρ twice.

Now suppose that ρ has capacity 4. In this situation, we have $\kappa \geq 1$, so we can apply the Core Copy Lemma to p/q and \widehat{p}/\widehat{q} .

Lemma 27.3 $\Sigma \subset R$.

Proof: The left edges of Σ and R both lie in the x -axis. So, we just have to show that the width of Σ is at most the width of R . The width of R is $\omega - 2\tau$. Statement 2 of Lemma 26.1 says that $\widehat{\kappa} = 0$. Hence, the width of $\widehat{\Sigma}$ is $\widehat{\tau}$, a quantity which not greater than $\widehat{\omega} - 2\widehat{\tau}$. (The latter quantity is the distance between the nearest line of capacity 4 to the x -axis.) Now, using Statement 5 of Lemma 26.1, we observe that $\widehat{\tau} \leq \widehat{\omega} - 2\widehat{\tau} = \omega - 2\tau$. ♠

Lemma 27.4 $\widehat{\Gamma}$ does not intersect the right edge of $\widehat{\Sigma}$.

Proof: Since $\widehat{\kappa} = 0$, the right edge of $\widehat{\Sigma}$ has capacity 2. $\widehat{\Gamma}$ only intersects the line containing this edge twice, and these intersection points must be outside of $\widehat{\Sigma}$, for otherwise $\widehat{\Gamma}$ could not make a closed loop. Compare Figure 2.1. ♠

It now follows from the Core Copy Lemma that Γ does not intersect the right edge of Σ . Figure 10.1 shows three sub-rectangles $R_1, R_2, \Sigma \subset R$. The horizontal dividers in the picture are the horizontal lines of capacity 2. We have already shown that Γ does not intersect the right edge of Σ . But then Γ cannot cross ρ between the two horizontal lines of capacity 2. To finish our proof, we just have to check that Γ intersects the right edge of R_1 once. By symmetry, Γ intersects the right edge of R_2 once, and the right edge of Σ blocks Γ so that it cannot intersect ρ anywhere else.

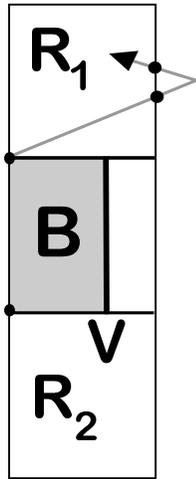


Figure 10.1: The rectangles R_1, S, R_2 .

We know that Γ cannot intersect ρ_1 more than twice because then, by symmetry, Γ would intersect ρ at least 6 times. If Γ intersects ρ_1 exactly twice, then Γ is trapped in R_1 and cannot get around to close up with the portion of Γ outside R_1 . Since Γ cannot get trapped in this way, we see that Γ intersects ρ_1 at most once. On the other hand, if Γ does not intersect ρ_1 at all, then Γ is trapped in the same way. In short, Γ intersects ρ_1 exactly once, and the same goes for ρ_2 .

This completes the proof of the Box Lemma.

27.4 Setup for the Copying Theorem

Let $R_{p/q}$ be the rectangle associated to the parameter p/q .

Lemma 27.5 *Let p'/q' be the even predecessor of p/q . Suppose that $\kappa = 0$. Then $R' \subset R$.*

Proof: Both boxes contain $(0, 0)$ as the bottom left vertex. So, we just have to show that the width and height of R' are at most the width and height of R . The width of R is τ . The width of R' is either τ' or $\omega' - 2\tau'$, whichever is smaller. In either case, we have $\text{width}(R') \leq \tau' \leq \tau = \text{width}(R)$. The height of R is ω and The height of R' is $\omega - 2\tau$. ♠

Suppose that p/q and p^*/q^* are two consecutive terms in the approximating sequence. We write

$$p/q = p_0/q_0 \prec p_1/q_1 \leftarrow \cdots \leftarrow p_n/q_n. \quad (289)$$

If p/q is strong then p_0/q_0 is the even predecessor of p_1/q_1 , and p_k/q_k is weak for $k = 1, \dots, n$. If p/q is core, then p_0/q_0 is the core predecessor of p_1/q_1 and p_1/q_1 is non-core, and p_k/q_k is weak for $k = 2, \dots, n$. Here $k \geq 2$.

We introduce some notation to help with the proof. Let $(\Sigma'_k, \Sigma_{k+1})$ be the pair of rectangles associated to the pair of parameters $(p_k/q_k, p_{k+1}/q_{k+1})$. Note that p_k/q_k gets two such rectangles attached to it, namely Σ_k and Σ'_k . These rectangles play different roles in the proof. Let Π_k denote the plaid tiling associated to p_k/q_k . In general, we set $\omega_k = \omega(p_k/q_k)$, etc.

27.5 The Strong Case

Suppose first that p_0/q_0 is strong. Recalling that $\Sigma'_0 = \Sigma_1 = B_0$, we conclude that

$$B_0 \cap \Pi_0 = \Sigma'_0 \cap \Pi_0 =^* \Sigma_1 \cap \Pi_1 = B_0 \cap \Pi_1. \quad (290)$$

The starred equality comes from the Strong Copy Lemma.

Recall that TH_k and BH_k are the top and bottom horizontal lines of capacity 2 with respect to p_k/q_k .

Lemma 27.6 *For each $k = 1, \dots, n$ the rectangle B_0 is contained in the lower half of B_k and one of BH_0 or TH_0 coincides with BH_k .*

Proof: Note that every box in sight contains $(0, 0)$ as the lower left vertices. So, we can decide which box contains which other box just by looking at the widths and heights.

Since $\tau_1 > 1/4$ we have $2\omega_0 < \omega_1$, the height of B_0 is less than half that of B_1 . Hence B_0 lies in the lower half of B_1 . By Lemma 27.5, we have $B_1 \subset \dots \subset B_k$. Hence B_0 lies in the lower half of B_k .

From Statement 1 of Lemma 26.1, we have either $BH_0 = BH_1$ or $TH_0 = BH_1$. From Statement 6 of Lemma 26.1, we see that $BH_1 = \dots = BH_k$. Hence either $BH_0 = BH_k$ or $TH_0 = BH_k$. ♠

Note that Σ'_k contains the lower half of B_k . Hence $B_0 \subset \Sigma'_k$ for all k . Now we will show inductively that $B_0 \cap \Pi_k$ implies $B_0 \cap \Pi_{k+1}$. We already have proved this for $k = 0$. For each $k = 1, \dots, n - 1$, we have $\tau_k < 1/4$. We have

$$B_0 \cap \Pi_0 = B_0 \cap \Pi_k \subset \Sigma'_k \cap \Pi_k =^* \Sigma_{k+1} \cap \Pi_{k+1}. \quad (291)$$

The starred equality comes from the Weak Copy Lemma. From this equation, we get $B_0 \cap \Pi_0 = B_0 \cap \Pi_{k+1}$. By induction, $B_0 \cap \Pi_0 = B_0 \cap \Pi_n$. That is, Π_n copies Π_0 inside B_0 , and B_0 lies in the lower half of B_n . This is what we wanted to prove.

27.6 The Core Case

Suppose that p_0/q_0 is core. Let Υ be the vertical translation by $(\omega_1 - \omega_2)/2$. By construction $\Upsilon(B_0)$ is symmetric with respect to the horizontal midline of B_1 . Statement 5 of Lemma 26.1 says that the distance between the two horizontal lines of capacity 2 is the same w.r.t p_0/q_0 and w.r.t. p_1/q_1 . Hence, by symmetry, $\Upsilon(TH_0) = TH_1$ and $\Upsilon(BH_0) = BH_1$. We have

$$\Upsilon(B_0 \cap \Pi_0) = \Upsilon(\Sigma'_0 \cap \Pi_0) = \Sigma_1 \cap \Pi_1 = \Upsilon(B_0) \cap \Pi_1. \quad (292)$$

Lemma 27.7 $\Upsilon(B_0) \subset \Sigma'_1$.

Proof: The left edge of $\Upsilon(B_0)$ lies in the y -axis, just like the left edge of Σ'_1 .

The width of $\Upsilon(B_0)$ is τ_0 and the width of Σ'_1 is $\omega_1 - 2\tau_1$. By Statement 2 of Lemma 26.1, the predecessor sequence cannot have 2 core terms in a row. Hence $\kappa_0 = 0$. This means that $3\tau_0 < \omega_0$. Hence

$$\tau_0 < \omega_0 - 2\tau_0 = \omega_1 - 2\tau_1.$$

The equality is Statement 5 of Lemma 26.1. This takes care of the widths.

The y -coordinate of the top edge of $\Upsilon(B_0)$ is $(\omega_0 + \omega_1)/2$. The height of Σ'_1 is either ω_1 or $2\tau_1$ depending on whether p_1/q_1 is strong or weak. Since $2\tau_1 < \omega_1$, we just need to deal with the weak case. That is, we have to show that $\omega_0 + \omega_1 < 4\tau_1$. But $4\omega_1/3 < 4\tau_1$ because p_0/q_0 is core. So, it suffices to show that

$$\omega_0 < \omega_1/3.$$

But this follows from Equation 272, because $\kappa_1 \geq 1$. ♠

Now we can take the next step in the argument.

$$\Upsilon(B_0 \cap \Pi_0) = \Upsilon(B_0) \cap \Pi_1 \subset \Sigma'_1 \cap \Pi_1 =^* \Sigma_2 \cap \Pi_2. \quad (293)$$

The starred equality is either the Weak Copy Lemma or the Strong Copy Lemma, whichever applies. Hence $\Upsilon(B_0 \cap \Pi_0) = \Upsilon(B_0) \cap \Pi_2$.

Lemma 27.8 *For each $k = 2, \dots, n$ the rectangle $\Upsilon(B_0)$ is contained in the lower half of B_k and one of BH_0 or TH_0 coincides with BH_k .*

Proof: By Lemma 27.5, the widths of R_1, \dots, R_k are non-decreasing. So, the width of $\Upsilon(B_0)$ is at most the width of R_k .

The y -coordinate of the top edge of $\Upsilon(B_0)$ is $(\omega_0 + \omega_1)/2$. The height of B_2 is ω_2 . By Statement 8 of Lemma 26.1, we see that the y coordinate of the top edge of $\Upsilon(B_0)$ is less than half the height of B_2 . This takes care of the case $k = 2$.

By Lemma 27.5, we have $B_2 \subset \dots \subset B_k$, and all these boxes have $(0, 0)$ as their bottom left vertex. Hence, $\Upsilon(B_0)$ lies in the bottom half of B_k .

We have already seen that $\Upsilon(BH_0) = BH_1$ and $\Upsilon(TH_0) = TH_1$. By Lemma 27.6, one of these two lines coincides with BH_2 . But then Statement 6 of Lemma 26.1 says that $BH_2 = \dots = BH_k$. ♠

The rest of the proof is the same inductive argument as in the Strong Case.

28 The Weak and Strong Copy Lemmas

28.1 The Mass and Capacity Sequences

In this chapter we prove the Weak and Strong Copy Lemmas. The two results have essentially the same proof, except for a few minor details.

Let $A = p/q$ be a parameter and let $\omega = p + q$ as usual. We also remind the reader that

$$P = \frac{2p}{\omega} = \frac{2A}{1+A}, \quad Q = \frac{2q}{\omega} = 2 - P. \quad (294)$$

In this section we repackage some of the results from §1.5. We fix integers $1 = x_0 < x_1 < \omega$ and $y_0 < y_1$. We define

$$\Sigma = [x_0, x_1] \times [y_0, y_1], \quad (295)$$

With this choice of x_0 and x_1 , the rectangle Σ does not intersect the vertical boundary of a block.

All the \mathcal{V} lines which intersect Σ have positive capacity. The signs could be positive or negative. We define two sequences $\{c_j\}$ and $\{m_j\}$ w.r.t. Σ . We have the *capacity sequence*

$$c_j = [2Pj]_2, \quad j = x_0, \dots, x_1. \quad (296)$$

The terms of the capacity sequence all lie in $(-1, 1)$.

We also have the *mass sequence*

$$m_j = [Pj + 1]_2, \quad j = y_0 - 2x + 1, \dots, y_1 + 2x - 1, \quad x = x_0 - x_1. \quad (297)$$

Remarks:

(i) Notice that the indices for the mass sequence start below the bottom edge of Σ (so to speak) and end above it. This is important for us for reasons which will become clear momentarily.

(ii) We will allow $j \in \omega\mathbf{Z}$ in the mass sequence. Such terms do not have a well-defined sign. The corresponding slanting lines are inert. When we speak of the signs of the terms of the mass sequence, we mean to ignore these terms.

(iii) Really we only care about the signs in the mass and capacity sequences, but for the purposes of running certain kinds of arguments it is useful to keep track of the numerical values as well.

Lemma 28.1 *The shade of any vertical intersection point in Σ is determined by the signs of the terms in the mass and capacity sequences.*

Proof: This is just a consequence of the equivalence of the two definitions of the plaid model. ♠

Corollary 28.2 *Suppose that we know how the plaid polygons intersect a single \mathcal{H} line inside Σ . Then the intersection of the plaid polygons with Σ is determined by the signs of the mass and capacity sequences.*

Proof: We will suppose that we know how the plaid polygons intersect the bottom edge of Σ . The case for any other edge has a similar treatment. Let Q be some unit square in Σ for which we have not yet determined the plaid model inside Q . We can take Q to be as low as possible. But then we know how the plaid model intersects the bottom edge of Q , and the signs of the mass and capacity sequences determine how the tiling intersects the left and right edges. But then the Fundamental Theorem for the plaid model tells us that ∂Q intersects the plaid polygons in either 0 or 2 points. This allows us to determine how the plaid polygons intersect the top edge of Q . ♠

28.2 A Matching Criterion

In this section, we mean to define all the objects in the previous section with respect to parameter p/q and p'/q' . Our notation convention will be that the object X corresponds to p/q whenever the same kind of object X' corresponds to p'/q' .

Let Π denote the union of plaid polygons with respect to p/q and let Π' denote the union of plaid polygons with respect to p'/q' . Suppose that Σ and Σ' are rectangles that are equivalent via a vertical translation Υ . That is, Υ preserves the y -axis and $\Upsilon(\Sigma') = \Sigma$. To be concrete, say that

$$\Upsilon(x, y) = (x, y + \xi). \tag{298}$$

We will give a criterion which guarantees that

$$\Upsilon(\Sigma' \cap \Pi') = \Sigma \cap \Pi. \tag{299}$$

Arithmetic Alignment: A necessary condition for Equation 299 is that the signs of $\{c_j\}$ are the same as the corresponding signs of $\{c'_j\}$ and the signs of $\{m_j\}$ are the same as the corresponding signs of $\{m'_j\}$. More precisely,

$$\text{sign}(c'_j) = \text{sign}(c_{j+\xi}), \quad \text{sign}(m'_j) = \text{sign}(m_{j+\xi}). \quad (300)$$

We say that (Σ, Π) and (Σ', Π') are *arithmetically aligned* if Equation 300 holds for all relevant indices.

It seems plausible that arithmetic alignment is sufficient for Equation 299 to hold, but we don't have a proof. We need more ingredients to make things work out cleanly.

Geometric Alignment: There is a natural correspondence between the vertical intersection points in Σ and the vertical intersection points in Σ' . Let z' be a vertical intersection point in Σ . Let $\{i', j'\}$ be the pair of indices so that the slanting lines through $(0, i')$ and $(0, j')$ contain z' . We let z denote the intersection of the slanting lines, of the same type, through $(0, i)$ and $(0, j)$. Here $i = i' + \xi$ and $j = j' + \xi$. We say that z' and z are *geometrically aligned* if z' and z are contained in the same unit vertical segment of Σ . We say that (Σ, Π) and (Σ', Π') are *geometrically aligned* if z and z' are geometrically aligned for every vertical intersection point $z' \in \Sigma'$.

It seems very likely that arithmetic and geometric alignment together imply Equation 299 but we don't have a proof. We need one more small ingredient.

Weak Horizontal Alignment: We say that (Σ, Π) and (Σ', Π') are *weakly horizontally aligned* if there are \mathcal{H} lines H' and H such that

$$\Upsilon(\Sigma \cap H \cap P) = \Sigma' \cap H' \cap P. \quad (301)$$

In other words, the tilings look the same on a single horizontal segment. This is exactly the criterion which appears in Corollary 28.2.

Now we come to our Matching Criterion.

Lemma 28.3 (Matching Criterion) *Suppose that*

- (Σ, Π) and (Σ', Π') are weakly horizontally aligned.
- (Σ, Π) and (Σ', Π') are geometrically aligned.

- (Σ, Π) and (Σ', Π') are arithmetically aligned.

Then $\Upsilon(\Sigma' \cap \Pi') = \Sigma \cap \Pi$.

Proof: Given Corollary 28.2, this is practically a tautology. The procedure given in Corollary 28.2 assigns exactly the same tiles to $\Sigma \cap \Pi$ as it does to $\Upsilon(\Sigma' \cap \Pi')$. ♠

28.3 Geometric Alignment

The idea is to verify the conditions of the Matching Lemma from §28.2. We use the notation and terminology from §28.2. Here we have $\Upsilon = \text{Identity}$ and hence $\Sigma = \Sigma'$.

The two plaid tilings agree along the bottom edge of Σ and Σ' , because this common edge lies in the boundary of the first block w.r.t. both parameters. Hence (Σ, Π) and (Σ', Π') are weakly horizontally aligned.

Lemma 28.4 (Σ, Π) and (Σ', Π') are geometrically aligned.

Proof: Let z and z' be corresponding points in $\Sigma = \Sigma'$. These points lie on slanting lines of the same type which have the same y -intercept. The difference in slopes of the two lines is

$$|P - P'| = |Q - Q'| = \frac{2}{\omega\omega'}.$$

Hence

$$\|z - z'\| \leq \frac{2\tau'}{\omega\omega'} < \frac{1}{\omega} < \frac{1}{\omega'} \quad (302)$$

But z' is at least $1/\omega'$ from the interval containing it. Hence z and z' lie in the same vertical unit interval. ♠

Remark: In replacing $1/\omega$ by $1/\omega'$, we threw away some of the strength of our estimate. We did this because in §9 we will have a much tighter estimate, and we want the two arguments to look similar.

28.4 Alignment of the Capacity Sequences

We introduce new variables

$$M_i = M_i\omega, \quad C_j = c_j\omega, \quad M'_i = m'_i\omega', \quad C'_j = c'_j\omega'. \quad (303)$$

- C_i is a nonzero even integer in $(-\omega, \omega)$.
- C'_i is a nonzero even integer in $(-\omega', \omega')$.
- M'_j is an odd integer in $[-\omega', \omega']$.
- M_j is an odd integer in $[-\omega, \omega]$.

Let

$$\lambda = \omega'/\omega. \quad (304)$$

We have

$$C'_i = [2\omega'P'i]_{2\omega'}, \quad C_i = [2\omega Pi]_{2\omega}, \quad \lambda C_i = [2\omega'Pi]_{2\omega'}. \quad (305)$$

We introduce the expression

$$\Psi(i) = |2\omega Pi - 2\omega'P'i|. \quad (306)$$

As long as

$$\Psi(i) < \ell + 2\lambda, \quad \min(|C'_i|, \omega' - |C'_i|) \geq \ell \quad (307)$$

the signs of C'_i and C_i are the same.

Using the fact that

$$|P - P'| = \frac{2}{\omega\omega'} \quad (308)$$

we see that

$$\Psi(i) = \frac{4i}{\omega} \leq \frac{4W'}{\omega}. \quad (309)$$

In all cases, we have $W' \leq \tau' \leq \omega'/2$. Hence

$$\Psi(i) < 2\lambda < 1 + 2\lambda.$$

So, we can take $\ell = 1$ in Equation 307. Hence, the two capacity sequences are arithmetically aligned.

28.5 Alignment of the Mass Sequences

We have

$$M'_i = [\omega' P' i + \omega']_{2\omega'}, \quad M_i = [2\omega P i + \omega]_{2\omega}, \quad \lambda M_i = [\omega' P i + \omega']_{2\omega'}. \quad (310)$$

We introduce the expression

$$W(i) = |\omega P i - \omega' P' i|. \quad (311)$$

This function differs from Ψ just in a factor of 2. As long as m'_i is a signed term and

$$W(i) < \ell + \lambda, \quad \min(|M'_i|, \omega' - |M'_i|) \geq \ell \quad (312)$$

the signs of M'_i and M_i are the same.

We have

$$W(i) = \frac{2i}{\omega} < \frac{2H' + 4W'}{\omega}. \quad (313)$$

There are several cases to consider.

Case 1: Suppose that we are in the weak case and $\tau' \leq 2\omega' - \tau'$. We have

$$H' = \omega' - \tau', \quad W' = \tau'.$$

Using the facts that

$$\tau' = \tau, \quad \omega' = \omega - 2\tau,$$

and plugging the values of W' and H' into Equation 313, we get

$$W(i) < \frac{2\omega' + 2\tau'}{\omega} = \frac{2\omega' + 2\tau}{\omega} = \frac{2\omega - 2\tau}{\omega} = \frac{\omega + \omega'}{\omega} = 1 + \lambda.$$

So, Equation 312 holds and there are no sign changes.

Case 2: Suppose that we are in the weak case and $\tau' > 2\omega' - \tau'$. This means that $\tau < \omega/4$ and $\tau' = \tau$ and $\tau' > \omega'/3$. We have

$$H' = 2\tau', \quad W' = \omega - 2\tau'.$$

Plugging this into Equation 313, we get

$$W(i) < \frac{4\omega' - 4\tau'}{\omega} < \frac{8\tau'}{\omega} = \frac{8\tau}{\omega} < 2.$$

So, we only have to worry about the case when $\widehat{C}_i = \pm 1$. This happens for $i \in \{-\tau', \tau', \omega' - \tau'\}$. The highest index is $\omega' + \tau' - 1$. we have

$$\begin{aligned} \Psi(\pm\tau') \leq \Psi(\omega' - \tau') &= \frac{2\omega' - 2\tau'}{\omega} = \frac{2\omega' - 2\tau}{\omega} = \frac{2\omega - 4\tau}{\omega} = \\ &= \frac{\omega + (\omega - 4\tau)}{\omega} < \frac{\omega + \omega'}{\omega} = 1 + \lambda. \end{aligned} \quad (314)$$

So Equation 312 holds in all cases.

Case 3: Suppose we are in the strong case and $\tau' \leq 2\omega' - \tau'$. We have

$$H' = \omega', \quad W' = \tau'.$$

This gives us the estimate

$$W(i) < \frac{2\omega' + 4\tau'}{\omega}. \quad (315)$$

Since $3\tau' \leq \omega'$ and $\tau + \tau' = \omega'$, we have

$$2\tau' \leq \omega' - \tau' = \tau.$$

Hence

$$W(i) < \frac{2\omega' + 2\tau}{\omega} = \frac{(\omega' + 2\tau) + \omega'}{\omega} = 1 + \lambda.$$

So, Equation 312 holds in all cases, and there are no sign changes.

Case 4: Suppose we are in the strong case and $\omega' - 2\tau' < \tau'$. In this case, we have

$$H' = \omega', \quad W' = \omega' - 2\tau'.$$

This gives us the estimate

$$W(i) < \frac{6\omega' - 8\tau'}{\omega}. \quad (316)$$

Using the fact that $\tau + \tau' = \omega'$ and $\tau < \omega/3$, we get

$$W(i) < \frac{6\tau - 2\tau'}{\omega} < 2.$$

Once again, we just have to worry about $C_i = \pm 1$. The relevant indices are $i \in \{-\tau', \tau', \omega' - \tau', \omega' + \tau'\}$. We have

$$\Psi(\pm\tau') \leq \Psi(\omega' - \tau') < \Psi(\omega' + \tau') = \frac{2\omega' + 2\tau'}{\omega} < \frac{\omega + \omega'}{\omega} = 1 + \lambda.$$

Here we used the fact that $2\omega' < \omega$ and $2\tau' < \omega'$. So, Equation 312 holds in all cases, and the mass sequences are aligned.

In short (Σ, Π) and (Σ', Π') are arithmetically aligned. We have verified all the conditions of the Matching Lemma, and so $\Sigma \cap \Pi = \Sigma' \cap \Pi'$. This proves the Weak and Strong Copy Lemmas.

29 The Core Copy Lemma

29.1 The Difficulty

We will again verify the criteria in the Matching Lemma from §28.2. We will change notation to reflect that we have been calling \widehat{p}/\widehat{q} the core predecessor of p/q

The general idea of the proof here is similar to what we did in the previous chapter, but here we must work with a weaker estimate, namely

$$|P - \widehat{P}| = |Q - \widehat{Q}| = \frac{4\kappa}{\omega\widehat{\omega}}. \quad (317)$$

The poorer quality of the estimate in Equation 317 forces us to work harder in certain spots of the proof.

29.2 Weak Horizontal Alignment

In our context here, recall that Υ is vertical translation by $\omega/2 - \widehat{\omega}/2$.

Lemma 29.1 $\Upsilon(0, \widehat{\tau}) = (0, \tau)$ and $\Upsilon(0, \widehat{\omega} - \widehat{\tau}) = (0, \omega - \tau)$.

Proof: Since Υ is vertical translation to $(\omega - \widehat{\omega})/2$, we have

$$\Upsilon(\widehat{\tau}) = \widehat{\tau} + (\omega - \widehat{\omega})/2 = \tau.$$

The last equality is Statement 5 of Lemma 26.1. A similar calculation shows that $\Upsilon(\widehat{\omega} - \widehat{\tau}) = \omega - \tau$. Here we have abused notation and just shown the action on the second coordinate. ♠

Lemma 29.2 (Σ, Π) and $(\widehat{\Sigma}, \widehat{\Pi})$ are weakly horizontally aligned.

Proof: We take \widehat{H} to be the line $y = \widehat{\tau}$ and we take H to be the line $y = \tau$. Since these lines have capacity 2 w.r.t. the relevant parameters, they only intersect the plaid tilings in the middle of the first interval. The second intersection points are outside $\widehat{\Sigma}$ and Σ respectively. ♠

29.3 Geometric Alignment

Lemma 29.3 (Σ, Π) and $(\widehat{\Sigma}, \widehat{\Pi})$ are geometrically aligned.

Proof: Let \widehat{z} and z be corresponding vertical intersection points. Let \widehat{U} and U be the two vertical intervals respectively containing \widehat{z} and z . We need to prove that $\Upsilon(\widehat{U}) = U$.

Let \widehat{L} and L be two slanting lines of the same type which contain \widehat{z} and z respectively. The two lines $\Upsilon(\widehat{L})$ and L have the same y intercept. The difference in their slopes is

$$|P - \widehat{P}| = |Q - \widehat{Q}| = \frac{4\kappa}{\omega\widehat{\omega}}. \quad (318)$$

Here $\kappa \geq 1$. These lines move at most $\widehat{\tau}$ in the horizontal direction. Hence

$$\|z - \Upsilon(\widehat{z})\| \leq \frac{4\kappa\widehat{\tau}}{\omega\widehat{\omega}}. \quad (319)$$

Statement 4 of Lemma 26.1 implies that $\widehat{\kappa} = 0$. Hence $\widehat{\tau} < \widehat{\omega}/3$. But then

$$\kappa\widehat{\tau} < \frac{1}{3}\kappa\widehat{\omega} <_* \frac{1}{3+2\kappa}\omega < \frac{\omega}{5}.$$

The starred inequality comes from Statement 7 of Lemma 26.1. Combining our equations, we get that

$$\|z - \Upsilon(\widehat{z})\| < \frac{1}{\widehat{\omega}}. \quad (320)$$

This means that both points z and $\Upsilon(\widehat{z})$ lie in the same vertical edge of a unit integer square. Hence $\Upsilon(\widehat{U}) = U$. ♠

29.4 Alignment of the Capacity Sequences

Define

$$\lambda = \widehat{\omega}/\omega, \quad \Psi(i) = |2\widehat{\omega}Pi - 2\widehat{\omega}\widehat{P}i|. \quad (321)$$

As long as Equation 307 holds, namely

$$\Psi(i) < \ell + 2\lambda, \quad \min(|\widehat{c}_i|, \omega - |\widehat{c}_i|) \geq \ell$$

the signs of \widehat{c}_i and C_i are the same.

Using Equation 317 and the fact that $i \in [1, \widehat{\tau}]$ we see that

$$\Psi(i) = \frac{8\kappa i}{\omega} \leq \frac{8\kappa \widehat{\tau}}{\omega} < \frac{8\kappa \widehat{\omega}}{3\omega} < \frac{8\kappa}{3+2\kappa} < 4. \quad (322)$$

As in the proof of Lemma 29.3 we have used the fact that $\widehat{\tau} \leq \widehat{\omega}/3$ and we have also used Statement 7 of Lemma 26.1. Now we know that $\Psi(i) < 4$. The only cases we need to worry about are

- $\widehat{C}_i = \pm 2$,
- $\widehat{C}_i = \pm(\widehat{\omega} - 1)$.
- $\widehat{C}_i = \pm(\widehat{\omega} - 3)$.

Case 1: We will consider the case when $\widehat{C}_i = 2$. The case when $\widehat{C}_i = -2$ has the same treatment. If $\widehat{C}_i = 2$ then $i = \widehat{\tau}$ and the corresponding \mathcal{V} line is the right edge of $\widehat{\Sigma}$. But then from Lemma 26.5 we know that the capacity of this line w.r.t is $4\kappa + 2$. So, we either have $C_i = 4\kappa + 2$ as desired or $C_i = -(4\kappa + 2)$. We will suppose that $C_i = -(4\kappa + 2)$ and derive a contradiction. In this case, we would have

$$\Psi(i) = 2 + \lambda(4\kappa + 2) = \frac{2\omega + 4\kappa\widehat{\omega} + 2\widehat{\omega}}{\omega}.$$

Combining this with Equation 322 we would get

$$2\omega + 4\kappa\widehat{\omega} + 2\widehat{\omega} < 8\kappa\widehat{\tau}.$$

Statement 4 of Lemma 26.1 says that $\widehat{\kappa} = 0$, which forces $\widehat{\tau} < \widehat{\omega}/3$, and this contradict the equation above.

Case 2: Lemma 2.1 tells us that the vertical lines of capacity $\pm(\widehat{\omega} - 1)$ either occur at $x = \pm\widehat{\tau}/2$ or $x = \pm(\widehat{\omega} - \widehat{\tau})/2, \text{ mod } \widehat{\omega}$. The latter case is irrelevant: The lines lie outside $\widehat{\Sigma}$. In the former case, the line of interest to is $x = \widehat{\tau}/2$. In other words $i = \widehat{\tau}/2$. Note that $2i = \widehat{\tau}$. The case $2i = \widehat{\tau}$ is the one just considered.

We will suppose that $\widehat{C}_{2i} = 2$. The case $\widehat{C}_{2i} = -2$ has the same treatment. When $\widehat{C}_{2i} = 2$, it means that

$$[\widehat{\omega}(2\widehat{P})(2i)]_{2\widehat{\omega}} = 2.$$

But this is the same as saying that

$$\widehat{\omega}(2\widehat{P})(2i) = 2\theta\widehat{\omega} + 2,$$

for some integer θ . But then

$$\widehat{\omega}(2\widehat{P}i) = 1 + \theta\widehat{\omega}.$$

Since the capacities are all even, this is only possible if θ is odd. But then

$$\widehat{C}_i = [\widehat{\omega}(2\widehat{P}i)]_{2\widehat{\omega}} = -\widehat{\omega} + 1.$$

So, \widehat{C}_i and \widehat{C}_{2i} have opposite signs. A similar argument shows that C_i and C_{2i} have opposite signs. Case 2 now follows from Case 1.

Case 3: When $\widehat{\tau}$ is even, the relevant line of capacity $\pm(\widehat{\omega} - 3)$ is the line $x = 3\widehat{\tau}/2$. Since $3\widehat{\tau}/2 < \widehat{\omega}/2$, this line is the one which intersects the first block and is closer to the y -axis. But $3\widehat{\tau}/2 > \widehat{\tau}$, and so our line does not intersect $\widehat{\Sigma}$.

When $\widehat{\tau}$ is odd, the relevant line of capacity $\pm(\widehat{\omega} - 3)$ is the line

$$x = \widehat{\omega}/2 - 3\widehat{\tau}/2.$$

It would happen that this line intersects $\widehat{\Sigma}$. We will deal with this line in a backhanded way.

We already know that $8\kappa\widehat{\tau}/\omega < 4$. If we really have a sign change in this case, we must have $\Psi(i) > 3$. But this means that

$$\frac{8\kappa\widehat{\tau}}{\omega} \in (3, 4). \tag{323}$$

we have

$$\Psi(i) = \frac{8\kappa i}{\omega} = \frac{4\kappa\widehat{\omega}}{\omega} - \frac{12\kappa\widehat{\tau}}{\omega}. \tag{324}$$

By Statement 7 of Lemma 26.1, the first term on the right hand side of this equation is at most 6. By Equation 323, the second term on the right lies in $(9/2, 6)$. Therefore $\Psi(i) \leq 3/2$, and we have a contradiction. Hence, there are no sign changes, and the two capacity sequences are arithmetically aligned.

29.5 Calculating some of the Masses

As a prelude to checking that the mass sequences are arithmetically aligned, we take care of some special cases. This will make the general argument easier.

We define the mass and sign of an integer point on the y -axis to be the mass and sign of the \mathcal{P} and \mathcal{Q} lines containing them. Unless explicitly stated otherwise, these quantities are taken with respect to the parameter p/q . In this section we establish some technical results about some of the masses and signs.

Lemma 29.4 *Let \widehat{p}/\widehat{q} be the core predecessor of p/q . Let Υ be the translation from the Core Copy Lemma.*

1. *Let Ω be the vertical interval of length $\widehat{\omega}$ and centered at the point $(0, \omega/2)$. The only points of mass less than $4\kappa + 2$ contained in Ω are the points of mass 1.*
2. *Υ maps the points of mass 1 w.r.t \widehat{p}/\widehat{q} to the points of mass 1 w.r.t p/q , and in a sign-preserving way.*
3. *Υ maps the points of mass $\widehat{\omega} - 2$ w.r.t \widehat{p}/\widehat{q} to the points of mass $\omega - 2$ w.r.t p/q , and in a sign-preserving way.*

We will prove Lemma 29.4 through a series of smaller steps. We first dispense with a minor technical point.

Lemma 29.5 $3\widehat{\tau} < \widehat{\omega}$.

Proof: Since \widehat{p}/\widehat{q} is assumed to be the nontrivial core predecessor of p/q , Statement 4 of Lemma 26.1 gives $\widehat{\kappa} = 0$. This forces $3\widehat{\tau} \leq \widehat{\omega}$. The case of equality would force $\widehat{p}/\widehat{q} = 1/2$. But then the even predecessor of p/q would be the even predecessor of $1/2$, by Statement 3 of Lemma 26.1. This is not possible. ♠

Proof of Statement 1: The points $(0, \tau)$ and $(0, \omega - \tau)$ are the two points of mass 1. We will give the proof when the sign of $(0, \tau)$ is positive. The other case is essentially the same.

Since Ω is symmetrically placed with respect to the horizontal midline of the first block $[0, \omega]^2$, it suffices to show that Ω does not contain any points of positive sign and mass $3, 5, 7, \dots, 4\kappa + 1$.

The endpoints of Ω are

$$(0, \omega/2 - \widehat{\omega}/2), \quad (0, \omega/2 + \widehat{\omega}/2).$$

Our proof refers to the work in §26.4. In particular, $h = \omega - 2\tau$ and $w = \widehat{\tau}$. Since the masses of the points on the y -axis occur in an arithmetic progression mod ω , the points on the y -axis having positive sign and mass $2\lambda + 1$, at least for $\lambda = 1, \dots, 2\kappa$, are

$$(0, \tau - \lambda h) + (0, \omega \mathbf{Z}). \quad (325)$$

The second summand is included just so that we remember that the whole assignment of masses is invariant under translation by $(0, \omega)$. Also, we say “at least” because once λ is large enough the sign will change. So, it suffices to prove that

$$\tau - \lambda h \in [-\omega/2 + \widehat{\omega}/2, \omega/2 - \widehat{\omega}/2], \quad \lambda = 1, \dots, 2\kappa.$$

Since these points occur in linear order, it suffices to prove the following two inequalities:

$$\tau - h < \omega/2 - \widehat{\omega}/2, \quad \tau - 2\kappa h > -\omega/2 + \widehat{\omega}/2. \quad (326)$$

Since $h = \omega - 2\tau$, the first identity is equivalent to $3\omega - 6\tau > \widehat{\omega}$. But, by Statement 5 of Lemma 26.1,

$$3\omega - 6\tau = 3\widehat{\omega} - 6\widehat{\tau}.$$

But the inequality

$$3\widehat{\omega} - 6\widehat{\tau} > \widehat{\omega}$$

is the same as the one proved in Lemma 29.5. This takes care of the first inequality.

It follows from Equation 280 in §26.4 that

$$\tau - 2\kappa h = -\omega + \tau + h + 2w.$$

Hence, the second inequality is the same as

$$(-\omega + \tau + h + 2w) - (-\omega/2 + \widehat{\omega}/2) > 0. \quad (327)$$

Plugging in $h = \omega - 2\tau$ and $w = \widehat{\tau}$, and using the relation $\omega - 2\tau = \widehat{\omega} - 2\widehat{\tau}$, we see that the left hand side of Equation 327 is just $\widehat{\tau}$. ♠

Remark: As a byproduct of our proof we note that the points $(0, \tau)$ and $(0, \tau + h + 2w)$ have the same sign. Also, $h + 2w = \widehat{\omega}$.

Proof of Statement 2: In view of Lemma 29.1, it suffices to prove that $(0, \widehat{\tau})$ has the same sign w.r.t. \widehat{p}/\widehat{q} as $(0, \tau)$ has w.r.t. p/q . We will consider the case when $(0, \widehat{\tau})$ has positive sign w.r.t \widehat{p}/\widehat{q} . The opposite case has essentially the same treatment.

The argument from Case 1 from §29.4 tells us that the horizontal lines through \widehat{y}_+ have capacity $4\kappa + 2$ w.r.t. p/q . Since this is twice an odd integer, Lemma 2.1 tells us that the slanting lines through \widehat{y}_+ have mass $2\kappa + 1$. Now we observe the following.

$$P\omega' = P'\omega' + (P - P')\omega' = 2p' \pm \frac{2}{\omega} \equiv \frac{2}{\omega} \pmod{2\mathbf{Z}}. \quad (328)$$

The calculation in Equation 328 implies that the masses w.r.t. p/q along the y -axis (so to speak) change by ± 2 when the y -coordinate changes by ω' . Since

$$\tau = \widehat{\tau} + k\omega'$$

we see that the slanting lines through τ either have mass 1 or have mass

$$4\kappa + 1 \pmod{\omega}.$$

But

$$\omega = \widehat{\omega} + 2\kappa\omega' > 6\kappa + 2.$$

Hence either the slanting lines through τ have mass 1 w.r.t p/q or they have mass $4\kappa + 1$. But we already know from Lemma 29.1 that these lines either have mass 1 or -1 . Since $4\kappa + 1 \neq -1$ we know that the slanting lines through τ have mass 1, as desired. ♠

Proof of Statement 3: Let z_{\pm} denote the points on the y -axis such that the slanting lines through these points have mass $\pm(\omega - 2)$. Likewise define \widehat{z}_{\pm} . We will treat the case when $(0, \tau)$ has positive sign. The other case has the same proof.

The function $f(x) \rightarrow [\widehat{P}x + 1]_2$ is locally affine, when the domain is interpreted to be the circle $\mathbf{R}/2\mathbf{Z}$. In the case at hand, this function changes

by $+2$ when we move from $\widehat{\tau}$ to $\widehat{\omega} - \widehat{\tau}$. Given this property, and the fact that $f(0) = \pm\widehat{\omega}$, we see that

$$f(\widehat{\omega} - 2\tau) = -\widehat{\omega} + 2.$$

But we also know that $y_+ = \tau$ and $y_- = \omega - \tau$ by the previous result. So, the same argument gives

$$f(\omega - 2\tau) = -\omega + 2.$$

This shows that $m_i = -\omega + 2$ when $\widehat{m}_i = -\widehat{\omega} + 2$. By the same argument, or symmetry, we see that $m_i = \omega - 2$ when $\widehat{m}_i = \widehat{\omega} - 2$. ♠

29.6 Alignment the Mass Sequences

We proceed as in §28.5. Let $\widehat{y} = \widehat{\omega}/2$ and $y = \omega/2$. Even though \widehat{y} is an integer point, note that

$$[\widehat{P}\widehat{y} + 1]_2 = [\widehat{p} + 1]_2, \quad [Py + 1]_2 = [p + 1]_2. \quad (329)$$

Since p and \widehat{p} are integers having the same parity, the two expressions above are the same. Hence, there is an integer θ so that

$$\widehat{\omega}\widehat{P}\widehat{y} = \widehat{\omega}Py + \theta\widehat{\omega}.$$

Recall that $i_0 = (\omega - \widehat{\omega})/2$ and $y = \widehat{y} + i_0$. We can re-write the last equation as $\Psi(\widehat{\omega}/2) = 0$, where

$$W(i) = |(\widehat{\omega}\widehat{P}i - \widehat{\omega}P(i + i_0) - \theta\widehat{\omega})| \quad (330)$$

The function W has the same properties as the similar function from Equation 312: As long as

$$W(i) < \ell + \lambda, \quad \min(|\widehat{M}_i|, \widehat{\omega} - |\widehat{M}_i|) \geq \ell$$

the signs of \widehat{M}_i and M_i are the same.

We say that the index i is *central* if $i \in (0, \widehat{\omega})$. Otherwise, we call i *peripheral*.

Lemma 29.6 *If i is a central index, and m'_i is term with a sign, then \widehat{m}_i and m_i have the same sign.*

Proof: When i is a central index, we have $|i - \widehat{\omega}/2| \leq \widehat{\omega}/2$. Combining this with Equation 317, we get

$$\Psi(i) \leq \frac{4\kappa}{\widehat{\omega}} \times \frac{\widehat{\omega}}{2} = \frac{2\kappa\widehat{\omega}}{\omega}. \quad (331)$$

Statement 7 of Lemma 26.1 says that $\kappa\widehat{\omega} < (3/2)\omega$. Combining this with the previous equation, we see that $\Psi(i) < 3$. This means that there are no sign changes unless $\widehat{m}_i = \pm 1$ or $\widehat{m}_i = \pm(\omega - 2)$. But Statements 2 and 3 of Lemma 29.4 take care of these special cases. ♠

Lemma 29.7 *If i is a peripheral index then \widehat{m}_i and m_i have the same sign.*

Proof: Note that all the peripheral indices are signed terms, because none of these indices are in $\widehat{\omega}\mathbf{Z}$.

The index j of each peripheral term m_j differs from the index i of some central term by $\pm\widehat{\omega}$. We set

$$j = i + \widehat{\omega}. \quad (332)$$

The case when $j = i - \widehat{\omega}$ has a similar treatment.

Consider the situation w.r.t. the parameter \widehat{p}/\widehat{q} first. These terms repeat every $\widehat{\omega}$. This means that $m'_j = m'_i$.

Now observe that

$$[P\widehat{\omega}]_2 = [\widehat{P}\widehat{\omega} + (P - \widehat{P})\omega]_2 = [(P - \widehat{P})\omega]_2 = \pm \frac{4\kappa}{\omega}. \quad (333)$$

Therefore

$$m_j = m_i \pm 4\kappa. \quad (334)$$

By Statement 1 of Lemma 29.4, the central terms w.r.t. p/q either equal ± 1 or are greater than $4\kappa + 2$. So, if $m_i \neq \pm 1$, we have the terms m_i and m_j have the same sign. But then

$$\sigma(m_j) = \sigma(m_i) = \sigma(m'_i) = \sigma(m'_j).$$

Here σ denotes the sign. The middle equality comes from the previous result about the central indices.

What about when $m'_i = \pm 1$? In this case (since $j = i + \widehat{\omega}$) we must have $i = \tau$ and $j = \tau + \widehat{\omega}$. The remark following the proof of Statement 1 of Lemma 29.4. takes care of this case. ♠

Now we know that (Σ, Π) and $(\widehat{\Sigma}, \widehat{\Pi})$ are arithmetically aligned. We have verified the conditions of the Matching Lemma, and so $\Sigma \cap \Pi = \Upsilon(\widehat{\Sigma} \cap \widehat{\Pi})$. This completes the proof of the Core Copy Lemma.

30 Existence of Many Large Orbits

30.1 The Empty Rectangle Lemma

In this chapter we prove Theorem 0.7. Fixing a parameter p/q and a block B and an even integer $K \geq 0$ let Γ_K denote the union of all the lines of capacity at most K which intersect B . The complement $B - \Gamma_K$ consists of $(K + 1)^2$ rectangles arranged in a grid pattern. We say that one of these rectangles is *empty* if its boundary has no light points on it. Empty rectangles serve as barriers, separating the plaid polygons inside them from the plaid polygons outside them.

Lemma 30.1 *For all parameters, all blocks B , and all choices of K , at least one of the rectangles of $B - \Gamma_K$ is empty.*

Proof: This is a counting argument. We will suppose that there are no empty boxes and derive a contradiction. There are a total of $(K + 1)^2$ rectangles. If some rectangle R has a light point on it, then it must have a second light point, because the polygon Γ crossing into R through an edge containing one of the light points must cross out of R through another edge.

The one exceptional situation is when the light point z lies at the corner of R . In this case, one of the edges E of R lies in a vertical boundary of the block B . Let's consider the case when E lies in the west boundary of R and z is the south west corner. The other cases are similar. Γ crosses into R through the south edge of R , but then it cannot exit through E because E lies in the boundary of B . So, even in this exceptional case, there must be 2 light points in the boundary of R .

If every rectangle has at least 2 light points, then there are at least $(K + 1)^2$ light points total. The idea is that we double the number of squares but then observe that we have counted some of the points twice.

On the other hand, we know that a line of capacity k contains at most k light points. Since there are 4 lines of capacity k for each $k = 0, 2, \dots, K$, this gives a total of

$$8 \sum_{k=1}^{K/2} k = (K + 1)^2 - 1.$$

We have one fewer point than we need. This is a contradiction. ♠

Remark: In practice, we see many empty rectangles within a block.

30.2 The Ubiquity Lemma

The result in this section follows from the Plaid Master Picture Theorem, the Unbounded Orbits Theorem, Compactness, and Equidistribution.

Lemma 30.2 (Ubiquity) *Let $\{p_k/q_k\} \subset (0, 1)$ be any sequence of even rational numbers with an irrational limit and let N be some fixed integer. Then there is some fixed number R , depending on the data above, with the following property. If $p_k \in B_k$ is any point then the disk of radius R about p_k intersects an N -fat polygon.*

Proof: Let $A = \lim p_k/q_k$. Let $P = 2A/(1 + A)$ and consider the PET $X_2(P)$ associated to the parameter A . Recall that $\Psi_A : G \rightarrow X_2(P)$ is the classifying map. Here G is the plaid grid.

By the Unbounded Orbits Theorem, there is some good offset $V \in \mathbf{R}^3$ so that one of the plaid polygons corresponding to $\Pi_{P,V}$ is fat. In other words, there is some point $\xi \in X_2(P)$ having an infinite orbit, such that the corresponding orbit in the plaid model is fat.

Recall that $X_2(P)$ is a subset of the affine PET X_2 . For any N , there is some $\epsilon > 0$ with the following property. If $\xi' \in X_2$ is less than ϵ from ξ , then the plaid polygon corresponding to ξ' is N -fat. The point here is that the orbit of ξ' converges to the orbit of ξ as $\epsilon \rightarrow 0$, and so the sequence of partition pieces containing the orbit of ξ' converges to the sequence of partition pieces containing the orbit of ξ . These sequences determine the plaid polygons up to translation.

At the same time, for any $\epsilon > 0$, we can take R sufficiently large so that the image

$$\Phi_A(G \cap B_R)$$

is $\epsilon/2$ -dense in $X_2(P)$. This follows from the fact that $X_A(G)$ is dense in $X_2(P)$, and from way that $X_A(G)$ equidistributes the torus $X_2(P)$. But then, for k sufficiently large, the image

$$\Phi_k(G \cap B_R)$$

is ϵ -dense in $X_2(P_k)$. Here we have set $\Phi_k = \Phi_{p_k/q_k}$ and $P_k = 2p_2/(p_k + q_k)$. So, for k sufficiently large, some point in $G \cap B_K$ contains a point μ such that

$$\|\Phi_k(\mu) - \xi\| < \epsilon.$$

By definition, the plaid polygon through μ is N -fat. ♠

30.3 Proof of the Main Result

Let N be as in Theorem 0.7. Say that a *fat disk* is a disk which satisfies the conclusions of the Ubiquity Lemma. This notion depends on N , but we fix N once and for all.

The Ubiquity Lemma says that every fat disk intersects an N -fat plaid polygon. When k is large there are many fat disks contained in B_k , but perhaps they all intersect the same N -fat plaid polygon. We want to use the Empty Rectangle Lemma to separate many of these fat disks so that the N -fat polygons they intersect are distinct. We do this by applying the Empty Rectangle Lemma many times, for different choices of the value K .

Assume for the moment that the parameter p/q is fixed. We call the grid Γ_K from the Empty Rectangle Lemma *fat* if it has sidelength at least $2N$. Suppose that $K < L$, we say that the two grids Γ_K and Γ_L are *totally different* if no rectangle of Γ_K is a rectangle of Γ_L . What we mean is that every rectangle of Γ_K is nontrivially subdivided into rectangles of Γ_L . If we apply the Empty Rectangle Lemma to Γ_K and Γ_L in this situation, we produce distinct empty rectangles which are either nested or have disjoint interiors. Following this section we will use some elementary number theory to establish the following result:

Lemma 30.3 (Grid Supply) *Let Ω be any positive integer. Once k is sufficiently large there are Ω pairwise totally different fat grids relative to p_k/q_k ,*

We choose $\Omega = 4^N$. It follows from the Ramsey Theorem that given Ω empty rectangles in the plane, which are either nested or disjoint, there are either N pairwise disjoint rectangles in the collection or N mutually nested rectangles in the collection.

In the disjoint case, we can find N pairwise disjoint fat disks D_1, \dots, D_N which are separated from each other by empty rectangles. We can also do this in the nested case, though it is less obvious. The idea in the nested case is that the region $R_1 - R_2$, where R_1 and R_2 are two nested empty rectangles, is a union of other rectangles all having minimum side length $2N$. Hence $R_1 - R_2$ contains a fat disk.

By the Ubiquity Lemma, we can find an N -fat plaid polygons P_j which intersects D_j . The polygons P_1, \dots, P_N are pairwise disjoint because they are all separated from each other by empty rectangles. This completes the proof of Theorem 0.7 modulo the proof of the Grid Supply Lemma.

30.4 The Continued Fraction Length

Given any number $x \in [0, 1]$, let $\Lambda(x)$ denote the length of the continued fraction of x . Note that $\Lambda(x)$ is finite if and only if x is rational. It is fairly easy to see that $\Lambda(p_k/q_k) \rightarrow \infty$ when p_k/q_k has an irrational limit.

However, the rational number which determines the geometry of the grids at the parameter p/q is $\widehat{\tau}/\omega$. Here $\omega = p+q$ and $\tau \in (0, \omega/2)$ satisfies $2p\widehat{\tau} \equiv 1 \pmod{\omega}$. The sequence $\{\widehat{\tau}_k/\omega_k\}$ might converge, and its subsequential limits might all be rational. Nonetheless, we will show that $\Lambda(\widehat{\tau}_k/\omega_k) \rightarrow \infty$. Once we know this, we will relate this fact to the geometry of circle rotations, and then we will prove the Grid Supply Lemma.

Note that the sequence $\{2p_k/\omega_k\}$ also has an irrational limit. Using this observation, the result we seek is an immediate consequence of the following result.

Lemma 30.4 *Suppose $\{A_n/C_n\}$ is an infinite sequence of rational numbers in $(0, 1)$ having an irrational limit. Suppose $A_n B_n \equiv 1 \pmod{C_n}$. Then $\Lambda(B_n/C_n) \rightarrow \infty$.*

Proof: Consider the two matrices

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \quad (335)$$

These matrices generate $PSL_2(\mathbf{Z})$. Here is a well-known estimate on $\Lambda(p/q)$. We find any other rational p'/q' such that $\det(M) = 1$, where

$$M = \begin{bmatrix} p & p' \\ q & q' \end{bmatrix}. \quad (336)$$

We write

$$M = A^{n_1} B^{n_2} \dots A^{n_k} \quad (337)$$

in the shortest possible way. Then $|\Lambda(p/q) - k|$ is uniformly bounded.

By definition, there is some integer K_n so that $A_n B_n - K_n C_n = 1$. But then $\Lambda(A_n/C_n)$ is comparable to the word length of

$$\begin{bmatrix} A_n & K_n \\ C_n & B_n \end{bmatrix} \quad (338)$$

Note that A and B are transposes of each other: $A = B^t$. But then the word length of M^t is the same as the word length of M . Finally, the word length of M^t is comparable to $\Lambda(B_n/C_n)$. ♠

30.5 Circle Rotations

Consider some $R = A/B \in (0, 1)$. We distribute K points x_0, \dots, x_{K-1} in the circle \mathbf{R}/\mathbf{Z} according to the formula

$$x_k = nR \pmod{\mathbf{Z}}. \quad (339)$$

The first K of these points x_0, \dots, x_{K-1} partition the circle into K intervals as long as $K < Q$. Call this partition $I'(K)$. Here we recall some well-known facts about the distribution of sizes of the intervals in $I'(K)$. There are either 2 different sizes of interval or 3 different sizes of interval. In case there are 3 different sizes, the sizes have the form A, B, C where $A + B = C$. When we add the next point to $I'(k)$ and produce $I'(k+1)$ the following two transitions can occur.

- If there are 3 interval sizes (A, B, C) then the new point divides the C interval into intervals of size A and B .
- If there are two interval sizes $A < B$, the new point divides a B -interval into two intervals of size A and $B - A$.

We say that the partitions $I'(K_1)$ and $I'(K_2)$ are *totally different* if the largest interval in $I'(K_2)$ is smaller than the smallest interval in $I'(K_1)$. We define $\Omega'(A, B)$ to be maximum number of totally different partitions.

Lemma 30.5 $\Omega'(A, B) > \Lambda(A/B)/6$.

Proof: Given a pair (A, B) of unequal positive integers, we define (A', B') to be whichever of $(A, B - A)$ or $(B, A - B)$ is positive. Starting with (A_0, B_0) we inductively define $(A_{j+1}, B_{j+1}) = (A'_j, B'_j)$ provided that $A_j \neq B_j$. We stop when $A_k = B_k$. In this case, the common value is the greatest common divisor of A_0 and B_0 . We call (A_i, B_j) and (A_j, B_j) *totally different* if $\min(A_i, B_i) > \max(A_j, B_j)$. We define the *Euclidean strength* of (A, B) to be maximum number of pairwise totally different pairs in the process. The strength of (A, B) is $\Lambda(A/B)/2$.

If we list out the gaps we see in our sequence of partitions, we see every pair that arises in the Euclidean algorithm. In each partition we see either 2 or 3 consecutive gaps from the total list. Hence, $\Omega(A, B)$ is at least $1/3$ the Euclidean strength of (A, B) . Hence $\Omega'(A, B) \geq \Lambda(A/B)/6$. ♠

30.6 Proof of the Grid Supply Lemma

Now we consider a modified process which is closer to the generation of the grids associated to the Grid Supply Lemma. We can distribute $2K$ points $x_0, y_0, \dots, x_{K-1}, y_{K-1}$ by setting x_j as above and $y_j = 1 - x_j$. The resulting partition $I(K)$ has $2K$ intervals. Notice that the distribution of gaps in $I(K)$ is the same as the distribution of gaps in $I'(2K - 1)$. We define $\Omega(A, B)$ just as we defined $\Omega'(A, B)$. The same argument as above shows that

$$\Omega(A, B) \geq \Lambda(A/B)/12. \tag{340}$$

Now we return to the notation from the Grid Supply Lemma. Combining Equation 340 with (the result deduced from) Lemma 30.4 we see that

$$\Omega(\widehat{\mathcal{T}}_k, \omega_k) \rightarrow \infty \tag{341}$$

as $k \rightarrow \infty$. We construct the grid Γ_K by distributing the coordinate-axes intercepts of the lines according to the modified circle rotation process. This shows that, once k is large enough, there are at least $\Omega 2 + N$ totally distinct grids. Since these grids are nested, the largest Ω of these grids will be N -fat. This completes the proof of the Grid Supply Lemma.

31 Tuned Sequences

31.1 Rescaling the Model

We say that $\{p_k/q_k\}$ is *tuned* if the sequence $\{\tau(p_k/q_k)\}$ also converges. We call $\lim \tau(p_k/q_k)$ the *tuned limit* of the sequence. By compactness, every sequence has a tuned subsequence. We call the tuned sequence *(ir)rationaly tuned* if the tuned limit is (ir)rational. In this chapter we prove the following result.

Theorem 31.1 *Let $\{p_k/q_k\} \subset (0, 1)$ be any irrationaly tuned sequence. Let $\{B_k\}$ be any sequence of associated blocks. Let N be any fixed integer. Then there is some $\delta > 0$ such that the following property holds once k is sufficiently large: More than N distinct plaid polygons have diameter more than $\delta\omega_k$, and every point of B_k is within ω_k/N of one of them.*

Let $P = 2A/(1 + A)$ and $Q = 2 - P$. There is a unique homothety T_k such that $T_k(B_k) = [0, 1]^2$. In this case, for each value of N , the rescaled grids

$$T_k(\Gamma_{k,N}) \tag{342}$$

converge (in the Hausdorff topology, say) to a grid G_M which has the following description.

- For each $k = 0, \dots, M$, the grid G_M contains the horizontal lines $y = k\tau$ and $y = 1 - k\tau$. These quantities are taken mod 1. These are rescaled limits of the horizontal lines of capacity k .
- For each $k = 0, \dots, M$, the grid G_M contains the vertical lines $x = k\tau$ and $x = 1 - k\tau$. These quantities are taken mod 1. These are rescaled limits of the vertical lines of capacity k .

Now we consider the \mathcal{P} lines and the \mathcal{Q} lines. Consider first the case when B_k is the first block for all k .

Lemma 31.2 *Let M be a positive odd integer. The rescaled limit $T_k(\Omega_{k,M})$ exists. It consists of the lines of slope $-P$ and $-Q$ which have y -intercept $\pm\mu\tau + \lambda$ for $\mu \in \{1, 3, 5, \dots, M\}$ and $\lambda \in \mathbf{Z}$.*

Proof: A calculation very much like the one in the proof of Lemma 2.1 shows that, for μ odd and $\lambda \in \mathbf{Z}$,

$$FP_k(0, \pm\mu a_k + \lambda\omega) = FP_k(0, \pm\mu) = \mp\mu/\omega_k. \quad (343)$$

Hence the \mathcal{P} and \mathcal{Q} lines through $(0, \pm\mu)$ have mass μ . Once k is sufficiently large, all these lines belong to $\Omega_{k,M}$ for $\mu = 1, 3, 5, \dots, M$. Moreover, no other points of the form $(0, y)$ satisfy $2py \equiv \pm\mu + \omega \pmod{2\omega}$. ♠

Now we consider the case when $\{B_k\}$ is an arbitrary sequence of blocks. Let B_k^0 denote the first block associated to p_k/q_k and let $\Omega_{k,M}^0$ denote the corresponding set of lines. The set $T_k(\Omega_{k,M}^0)$ differs from the set $T_k(\Omega_{k,M})$ by a vertical translation. We can take this translation to have length at most 1 because our sets are both invariant under vertical translation by 1. So, by compactness, we can pass to a subsequence and assume that $T_k(\Omega_{k,M}^0)$ really does converge. The limit is the set of lines of slope $-P$ and $-Q$ having y -intercept $\pm\mu + \lambda + \xi$, where $\mu \in \{1, 3, 5, \dots, M\}$ and $\lambda \in \mathbf{Z}$ and $\xi \in (0, 1)$ is the translation factor. We call ξ the *offset* of the limit.

So, if we pass to a subsequence, then the sets $\{T_k(\Gamma_{k,M})\}$ and $\{T_k(\Omega_{k,M})\}$ converge to Γ_M and Ω_M respectively. We assign a capacity to the lines in Γ_M in the obvious way: If some line is the limit of lines of capacity c , it gets capacity c . Likewise, we assign a mass to the lines in Ω_M . This allows us to assign a set of light Σ_M on the lines of Γ_M .

By construction, the sets $T_k(\Sigma_{k,M})$ *practically converge* to the set Σ_M . There is one case we have to worry about. If Σ_M contains a point in the corner of $[0, 1]^2$ then it might not arise as the limit of points in $T_k(\Sigma_{k,M})$. This situation would not happen if B_k is always the first block, but it could happen in general. But we can say that every point of $\Sigma_M \cap (0, 1)^2$ is the limit of points in $T_k(\Sigma_{k,M})$. Moreover, if a light point in $\Sigma_M \cap (0, 1)^2$ is at least δ from every other point of Σ_M , then the corresponding point of $\Sigma_{k,M}$ is separated by $\delta\omega_k$ from every other light point in $\Sigma_{k,M}$.

31.2 The Filling Property

Now we recall a familiar fact about circle rotations. Let $\theta \in (0, 1)$ be irrational. Consider the map

$$T(x) = x + \theta \pmod{1}. \quad (344)$$

For any $\epsilon > 0$ there is some M such that the first M points of the orbit $\{T^j(x)\}$ is ϵ -dense. The value of M only depends on θ and ϵ and not on the starting point x . The way that M depends on θ and ϵ is subtle; it has to do with the continued fraction expansion of θ . However, we do not care about this subtlety.

Here is a consequence of the filling property. We keep the notation from the previous section.

Lemma 31.3 *Let D be a disk of radius ϵ in $[0, 1]^2$. Then there is a constant M and some horizontal line L of Γ_M so that $L \cap D \cap \Sigma_M$ contains at least 2 light points.*

Proof: We say that a line L *frankly intersects* a disk D if $L \cap D$ contains a point which is within $\text{radius}(D)/4$ from the center of D . If a horizontal line and a line of slope $-Q \in (-1, -2)$ both frankly intersect D , then their intersection is contained in D .

By the filling property, there is some M' such that at least 2 lines of $\Omega_{M'}$ frankly intersect D . Call these lines Q_1 and Q_2 . We can take these lines to be of type \mathcal{Q} and of positive sign. Also by the filling property, there is some $M > M'$ so that at least one horizontal line L having positive sign and capacity in (M', M) frankly intersects D .

We get out two light points of $L \cap D \cap \Sigma_M$ by intersecting L with the two lines Q_1 and Q_2 . ♠

We say a bit more about the light points produced by the previous lemma.

Lemma 31.4 *Suppose that z_1 and z_2 are two points of $\Sigma_M \cap D \cap \Gamma_M$ that lie on the same horizontal line. Then at least one of the two points is distinct from every other point of $\Sigma_M \cap L$.*

Proof: Let L be the line $y = \xi_1$. Let ξ_2 be the offset of our limit. Our points have the form

$$z_j = \left(\frac{\mu_j \tau + \xi_3}{Q} \right), \quad |\mu| \leq M, \quad \xi_3 = \xi_1 + \xi_2. \quad (345)$$

Given the irrationality of τ , these two points are distinct from each other, and also distinct from every other point of type \mathcal{Q} on L .

Suppose then that z_1 and z_2 are both points of type P as well. Then we have

$$|z_1 - z_2| = \frac{c_1}{P} = \frac{c_2}{Q}, \quad c_1, c_2 \in \mathbf{Z}. \quad (346)$$

This contradicts the fact that P/Q is irrational. ♠

31.3 Proof of Theorem 31.1

Our proof is similar in spirit to what we did for the proof of Theorem 0.7, except that we have more control over the diameters of the big plaid polygons involved. We will work with the rescaled limit and then, at the end, interpret what our result says.

The Empty Rectangle Lemma applies to the grid Γ_M . By varying M and applying the Empty Rectangle Lemma N times, we can find N rectangles in $[0, 1]^2$, say R_1, \dots, R_N such that each R_j contains a disk D which is disjoint from R_i for $j < i$. The filling property lets us take N as large as we like.

By the work in the previous section, each disk D_j will contain a light point z_j , which is distinct from all other light points on the same horizontal line. Let 2δ be the minimum separation between z_j and any other light point. The minimum is taken over all $j = 1, \dots, N$. By making δ smaller, if necessary, we can arrange that every disk of radius $1/2N$ contains the kind of light point which is separated from its horizontal neighbors by at least 2δ .

Now let us see what this says about the picture in the block B_k . Once k is sufficiently large, we can find rectangles $R_{k,1}, \dots, R_{k,N}$ which contain disks $D_{k,1}, \dots, D_{k,N}$ having the following properties.

- $D_{k,j}$ is disjoint from $R_{k,i}$ if $j < i$.
- $D_{k,j}$ contains a light point $z_{k,j}$ which is separated from all other light points on the same horizontal line by at least $\delta\omega_k$.

Let $\Gamma_{k,j}$ denote the plaid polygon that intersects the horizontal unit segment containing z_j . Since $\Gamma_{k,j}$ is a closed loop, it must intersect the horizontal line containing $z_{k,j}$ in a second light point. Hence $\Gamma_{k,j}$ has x -diameter at least $\delta\omega_k$. By construction, the polygons $\Gamma_{k,j}$ and $\Gamma_{k,i}$ lie in different components of $B_k - \partial R_{k,j}$ for $j < i$. Hence, these polygons are all distinct.

At the same time, choose any tile center $c \in B_k$. Let Δ be the disk of radius ω_k/N about c . Then $T_k(\Delta)$ contains a disk of radius $1/2N$ for k

large. Hence $T_k(\Delta)$ contains a light point that is separated from its horizontal neighbors by at least 2δ . But then the inverse image of this point is a light point in Δ that is separated from all the other light points on the same horizontal line by at least $\delta\omega_k$. Hence, c is within ω_k/N of some plaid polygon having diameter at least $\delta\omega_n$.

This completes the proof of Theorem 31.1.

32 References

- [DeB] N. E. J. De Bruijn, *Algebraic theory of Penrose's nonperiodic tilings*, Nederl. Akad. Wetensch. Proc. **84**:39–66 (1981).
- [DF] D. Dolyopyat and B. Fayad, *Unbounded orbits for semicircular outer billiards*, Annales Henri Poincaré **10** (2009) pp 357-375
- [G] D. Genin, *Regular and Chaotic Dynamics of Outer Billiards*, Pennsylvania State University Ph.D. thesis, State College (2005).
- [GS] E. Gutkin and N. Simanyi, *Dual polygonal billiard and necklace dynamics*, Comm. Math. Phys. **143**:431–450 (1991).
- [H] W. Hooper, *Renormalization of Polygon Exchange Transformations arising from Corner Percolation*, Invent. Math. **191.2** (2013) pp 255-320
- [Ko] Kolodziej, *The antibilliard outside a polygon*, Bull. Pol. Acad. Sci. Math. **37**:163–168 (1994).
- [M] J. Moser, *Is the solar system stable?*, Math. Intelligencer **1**:65–71 (1978).
- [N] B. H. Neumann, *Sharing ham and eggs*, Summary of a Manchester Mathematics Colloquium, 25 Jan 1959, published in Iota, the Manchester University Mathematics Students' Journal.
- [S1] R. E. Schwartz, *Outer Billiard on Kites*, Annals of Math Studies **171** (2009)
- [S2] R. E. Schwartz, *Outer Billiards and the Pinwheel Map*, J. Mod. Dyn. **2**:255-283 (2011).
- [S3] R. E. Schwartz, *Outer Billiards, Polytope Exchange Transformations, and Quarter Turn Compositions*, preprint (2013)
- [T1] S. Tabachnikov, *Geometry and billiards*, Student Mathematical Library 30, Amer. Math. Soc. (2005).
- [T2] S. Tabachnikov, *Billiards*, Société Mathématique de France, “Panoramas et Synthèses” 1, 1995
- [VS] F. Vivaldi and A. Shaidenko, *Global stability of a class of discontinuous dual billiards*, Comm. Math. Phys. **110**:625–640 (1987).