Geometric and algebraic properties of polyomino tilings

by

Michael Robert Korn

B.A., Princeton University, 2000

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2004

Author . . . .
Department of Mathematics
May 4, 2004

Certified by . . . .
Igor Pak
Assistant Professor of Mathematics
Thesis Supervisor

Accepted by . . . .
Rodolfo Ruben Rosales
Chairman, Committee on Applied Mathematics

Accepted by . . . .
Pavel Etingof
Chairman, Department Committee on Graduate Students

Geometric and algebraic properties of polyomino tilings

by

Michael Robert Korn

Submitted to the Department of Mathematics
on May 4, 2004, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In this thesis we study tilings of regions on the square grid by polyominoes. A polyomino is any connected shape formed from a union of grid cells, and a tiling of a region is a collection of polyominoes lying in the region such that each square is covered exactly once. In particular, we focus on two main themes: local connectivity and tile invariants.

Given a set of tiles $\mathcal{T}$ and a finite set $\mathcal{L}$ of local replacement moves, we say that a region Γ has local connectivity with respect to $\mathcal{T}$ and $\mathcal{L}$ if it is possible to convert any tiling of Γ into any other by means of these moves. If $\mathcal{R}$ is a set of regions (such as the set of all simply connected regions), then we say there is a local move property for $\mathcal{T}$ and $\mathcal{R}$ if there exists a finite set of moves $\mathcal{L}$ such that every Γ in $\mathcal{R}$ has local connectivity with respect to $\mathcal{T}$ and $\mathcal{L}$. We use height function techniques to prove local move properties for several new tile sets. In addition, we provide explicit counterexamples to show the absence of a local move property for a number of tile sets where local move properties were conjectured to hold.

We also provide several new results concerning tile invariants. If we let $a_i(\tau)$ denote the number of occurrences of the tile t_i in a tiling τ of a region Γ, then a tile invariant is a linear combination of the a_i's whose value depends only on Γ and not on τ. We modify the boundary-word technique of Conway and Lagarias to prove tile invariants for several new sets of tiles and provide specific examples to show that the invariants we obtain are the best possible.

In addition, we prove some new enumerative results, relating certain tiling problems to Baxter permutations, the Tutte polynomial, and alternating-sign matrices.

Thesis Supervisor: Igor Pak
Title: Assistant Professor of Mathematics

Acknowledgments

I would first and foremost like to thank my advisor, Professor Igor Pak, for suggesting interesting problems, keeping me on track, and putting up with my crazy American attitudes about work. His knowledge of the subject and cheerful demeanor were greatly appreciated during the course of writing this thesis.

Hearty thanks also go to Professor Daniel Kleitman and Professor Richard Stanley for being part of my thesis committee, providing helpful comments, and being very supportive of my studies. They have been excellent role models throughout my time at MIT.

Thanks should also go to the Department of Defense and the American Society for Engineering Education, for providing me with a generous fellowship which made my graduate studies, and ultimately this thesis, possible.

Furthermore, I would like to thank my officemates Andrew Brooke-Taylor, Rados Radoicic, Karen Bernhardt, and Frederic Rochon, the entire MIT math department, as well as Monica Linden, Carol Meyers, Ben Walter, and everyone else I've known here, for making this experience a pleasant one.

Perhaps most importantly, I thank my parents for giving me reason to believe in myself.

Contents

List of Figures

Chapter 1

Introduction

1.1 Polyominoes

The problems we will be considering in this work take place on the unit grid in two dimensions. The unit grid is composed of unit squares with integer coordinates, which we will call *cells*.

A *polyomino* is any connected shape formed by attaching together cells of the unit grid. We will require that cells be connected by edges, not merely at their corners. We will also allow polyominoes to have interior holes (although some other authors only use the term “polyomino” for simply-connected shapes). Figure 1-1 shows some polyominoes, while Figure 1-2 shows some shapes which are not polyominoes.

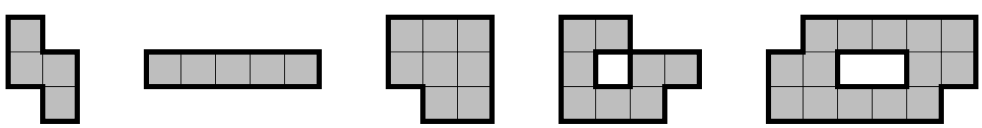


Figure 1-1: Some polyominoes.

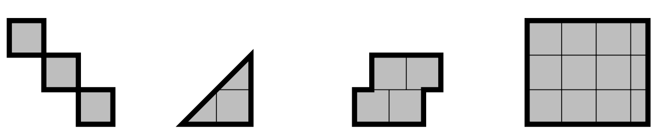


Figure 1-2: Some shapes which are not polyominoes.

The word “polyomino” was coined by Solomon Golomb in 1953, in his book, *Polyominoes* [13]. The shape formed by two unit squares placed edge to edge is called a domino, because of its resemblance to the pieces from the game of dominoes. Golomb then dropped the "d", and used the suffix "-omino" to describe shapes formed by attaching any number of squares together in this way. For instance, a *tromino* (also called *triomino*) is formed by attaching three squares together (either in a line or an L-shape). Four cells together make a *tetromino*, while five together are called a *pentomino*. A *monomino* is just one unit square by itself. *Polyominoes* are just *-ominoes* of any size.

When are two polyominoes the same? It depends on whether we consider them to be free, one-sided, or fixed. *Free* polyominoes are allowed to be rotated and reflected. *One-sided* polyominoes are allowed to be rotated but not reflected. *Fixed* polyominoes cannot be rotated or reflected. (Translation is always allowed.) For instance, the computer game Tetris is played using the seven one-sided tetrominoes (because the shapes can be rotated, but not reflected). For our purposes, we will always consider polyominoes to be fixed.

1.2 Tilings

Let Γ be a region of the plane. All we mean by this is that Γ is a polyomino, but we use the word "region" instead to avoid confusion. A *tiling* of Γ is a collection of polyominoes situated inside Γ such that each cell of Γ is covered by exactly one of these polyominoes. We call the individual polyominoes *tiles*.

Here is a well-known tiling "puzzle" problem:

> There are twelve distinct free pentominoes (Figure 1-3). Can you arrange these twelve pieces to form a 6×10 rectangle?

There are a number of other puzzles like this, where you are given a specified collection of polyominoes and asked to arrange them into a particular shape. However, the tiling problems that we will be most interested in will be of a slightly different sort.

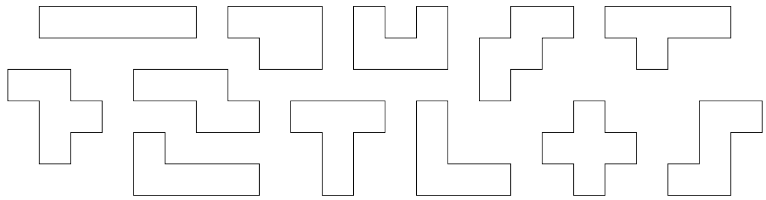


Figure 1-3: The twelve free pentominoes.

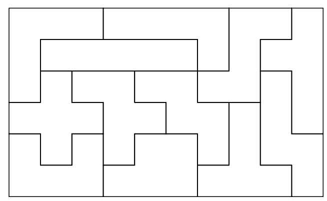


Figure 1-4: A tiling of a 6×10 rectangle using the twelve free pentominoes.

Generally we will consider problems where we are given a finite region Γ to be tiled, and a set of polyominoes $\mathcal{T}$, and we will be concerned with tilings of Γ such that each tile is equivalent to some shape in $\mathcal{T}$. In contrast to the previous example, here we will allow ourselves to use tiles as many times as we need, or possibly not at all. Also, as we said before, we will restrict ourselves to fixed polyominoes—we are not allowed to rotate or reflect the shapes in $\mathcal{T}$. This actually gives us greater flexibility in designing problems. If we wish to allow rotations, we may just expand $\mathcal{T}$ to include each orientation of each shape separately.

The following classical "brain-teaser" is very similar to the types of tiling problems we will be interested in.

> Consider a standard 8×8 checkerboard. Suppose you have a suppy of dominoes, each of which can cover exactly two squares of the checkerboard. Notice that with 32 dominoes, you can cover the entire board (Figure 1-5 shows one of the many ways of doing this). Now consider a modified checkerboard, where we cut off two squares on opposite corners of the board. Can you now cover the modified board with 31 dominoes?

The answer, of course, is no. Each domino always covers one red square and one black square. However, the modified checkerboard contains 32 red squares and 30 black squares, so it cannot be covered by 31 dominoes.

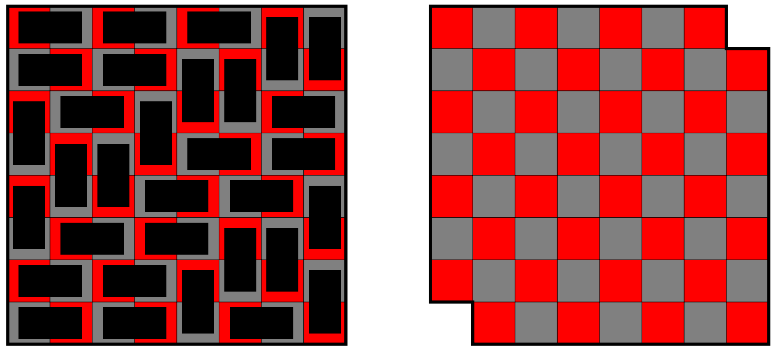


Figure 1-5: A regular checkerboard tiled with dominoes, and our modified checkerboard.

In the next two chapters, we will see two aspects of tilings that will be of particular interest to us: local moves and tile invariants. Chapter 2 provides an overview of local moves. We give definitions and some motivation, as well as a review of the basic techniques that can be used to prove (and disprove) local move properties—techniques that we will use later in this paper. In a similar fashion, chapter 3 describes the notion of tile invariants. Again we provide definitions and a few examples, and we explain the general techniques that we will use to prove tile invariants later in the paper.

The remaining chapters can be read independently of each other. In chapter 4 we investigate the set of T-tetrominoes. There is an interesting structure which emerges from such tilings, allowing us to prove a local move property as well as several enumerative results. Chapter 5 deals with tiling problems where either the tiles, or the region to be tiled, are rectangles. We will see specific examples where local move properties hold, and examples where they do not hold. In chapter 6, we again deal with rectangles, but this time we consider rectangles whose side lengths are irrational. This setup leads to an unexpected connection to a class of permutations called Baxter permutations. In chapter 7, we define a new set of tiles which we call the *generalized dominoes of order k*. We will show that the generalized dominoes have a very nice local move property (which generalizes the local move property for ordinary dominoes). In chapter 8, we consider two sets of tiles, first the horizontal T-tetrominoes and horizontal skew-tetrominoes, and second the horizontal T-tetrominoes and horizontal domino. In the latter case, we are able to prove a local move property, while in the former case no local move property holds, but there is an unexpected tile invariant which appears. Chapter 9 deals with the set of skew-tetrominoes. Here there is no local move property, but again we are able to prove a non-trivial tile invariant. In chapter 10, we prove tile invariants for the so-called ribbon tiles. The results in this chapter were known before, but we supply a more elegant proof, based upon the previous proof idea. In chapter 11, we consider the problem of building a tileable region from an untileable region by adding tiles to the outside of the region. We show that there is a class of regions for which this is always possible, but that the result in not true in general. Finally, in chapter 12 we provide a specific set of three-dimensional tiles having a specific property which was asked for in [18].

Chapter 2

Local moves

2.1 Introduction

Given a set of tiles $\mathcal{T}$ and a region Γ, let $\mathcal{Y}$ denote the set of all possible tilings of this region with these tiles. We would like to know something about the structure of the set $\mathcal{Y}$. Some tilings may be very similar, only differing in a small area, while others may look completely different. In order to capture this idea, we consider the notion of a *local move*.

A local move is an operation which converts one tiling of a region into another. Essentially, a local move consists of "picking up" some number of tiles from a tiling, then placing new tiles to fill that space in a different way. For instance, suppose we have a region tiled by dominoes. If there are two vertical dominoes next to each other forming a 2×2 square, then we may pick up these tiles and replace them with two horizontal tiles filling that same 2×2 square (Figure 2-1).

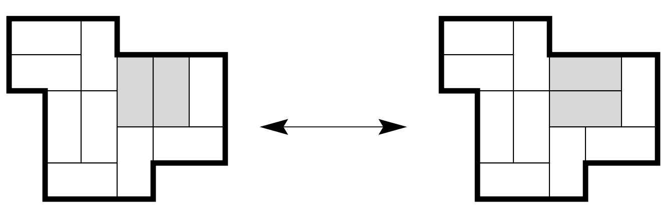


Figure 2-1: A local move applied to a domino tiling.

With no other restrictions, anything could be a local move—we could pick up all the tiles and then place new tiles to create a completely different tiling, and call this a local move. To avoid this situation, we usually specify a particular finite collection of allowable moves. An allowable move is a small region which can be tiled in two or more ways. If we have a large tiling where some subset of its tiles forms this smaller region, then we are allowed to pick up those tiles, and refill the space with any of the other possible tilings of that small region.

Suppose $\mathcal{T}$ is a set of tiles. Let $\mathcal{L}$ be a finite set of allowable local moves for this tile set. So $\mathcal{L}$ is a set of (small) regions, each of which can be tiled in at least two ways. Let Γ be any tileable region. Define the *local move graph* for $(\mathcal{T}, \mathcal{L}, \Gamma)$ by placing a vertex for each tiling of Γ, and placing an edge between two vertices if the corresponding tilings are related by a local move which belongs to $\mathcal{L}$. If this graph is connected, we will say that Γ has *local connectivity* with respect to $\mathcal{L}$ (and $\mathcal{T}$). (We may also say that Γ is *locally connected* with respect to $\mathcal{L}$ and $\mathcal{T}$.) We will say two tilings are *local-move equivalent* if they lie in the same connected component of this graph.

Local connectivity is interesting for several reasons. Notice that this local move graph allows us to define a Markov chain on $\mathcal{Y}$. (See [1] for basic definions and properties of Markov chains.) Specifically, we consider the Markov chain where at each step, we are allowed to move from one vertex of the local move graph to an adjacent one. If Γ has local connectivity, then the graph is connected, so the Markov chain is connected as well. If we set up the transition probabilities to be all the same, and also allow the possibility of remaining at the same vertex in a time step, then the Markov chain will be symmetric, ergodic, and connected. Hence the Markov chain will converge upon the uniform distribution, so running the Markov process long enough will allow us to sample tilings from $\mathcal{Y}$ uniformly at random.

Another reason that local connectivity is interesting is that it allows us to prove invariants. Suppose Ψ is some property of a tiling, and Ψ is preserved upon applying any local move from $\mathcal{L}$. Then if Ψ holds for one tiling of Γ, and Γ has local connectivity, then Ψ must hold for all tilings of Γ. For instance, consider domino tilings, and let Ψ be the property that the number of vertical dominoes in a tiling of Γ is odd. If every local move in $\mathcal{L}$ preserves the parity of the number of vertical dominoes, and one tiling of Γ contains an odd number of vertical dominoes, and Γ has local connectivity with respect to $\mathcal{L}$, then every tiling of Γ must use an odd number of vertical dominoes. The only drawback to this approach is that local connectivity is usually very difficult to prove compared to other methods of proving such invariants.

2.2 Local connectivity for sets of regions

While it is interesting to know that some region Γ has local connectivity with respect to a set of local moves, it would be more interesting if this set of local moves worked for a lot of regions. Let $\mathcal{R}$ be a set of regions. Some common examples would be the set of all regions, all simply-connected regions, or perhaps the set of all rectangles. Let us denote these by $\mathcal{R}_{all}$, $\mathcal{R}_{sc}$, and $\mathcal{R}_{rect}$ respectively.

We say that there is a *local move property* for $\mathcal{T}$ and $\mathcal{R}$ if there exists a set of local moves $\mathcal{L}$ such that every region Γ in $\mathcal{R}$ has local connectivity with respect to $\mathcal{L}$.

2.2.1 The set of all regions

The best situation, it seems, would be a tile set which has a local move property for all regions Γ. Unfortunately, this seems never to be the case except for tile sets $\mathcal{T}$ which are trivial in some way.

For example, take $\mathcal{T}$ to be the set of tiles shown in Figure 2-2. These tiles are so restrictive that there are no regions which admit more than one tiling. (An algorithm for tiling is as follows: Consider the left-most cell in the top row of Γ. If the cell to its right belongs to Γ, place the 2×3 hexomino there. Otherwise place the U-pentomino there. Continue with the portion of Γ that remains untiled.) Hence the local move graph contains at most one vertex, and thus is always connected (even though we have no local moves). This is not interesting.

Another example is if we take $\mathcal{T}$ to be any tile set which includes the monomino (the 1×1 square). We may define a set of local moves as follows. For each tile, take a region of that shape. This region may be tiled with the one tile of that shape, or it

Figure 2-2: A set of tiles which can tile any region in at most one way. Rotations are not allowed.

may be tiled by monominoes. Take any region Γ. Let τ_0 be the tiling of Γ which uses only monominoes. Let τ be any tiling. If τ contains a tile which is not a monomino, then we may perform a local move and replace it with monominoes. In this way, we may perform local moves on τ until we reach τ_0. Thus every tiling is in the same component of the local move graph as τ_0, so the local move graph is connected. This result is somewhat more interesting than the first, but we will still classify it as a "trivial" case.

We do not know of any tile set $\mathcal{T}$ which has a local move property for all regions that does not rely on a trick similar to the ones above. It would be very interesting to see a truly non-trivial set of tiles which has this property.

In section 2.4 we will see why many tile sets cannot have a local move property for the set of all regions.

2.2.2 The set of simply-connected regions

If we only want local connectivity for all simply connected regions, then there are a number of tile sets $\mathcal{T}$ for which this is possible. The most famous of these is the set of dominoes. With dominoes, a 2×2 square can be tiled in two ways. It turns out that this local move is sufficient to give local connectivity for all regions in $\mathcal{R}_{sc}$. So if Γ is any simply-connected region, then one may convert any domino tiling of Γ into any other by flipping two dominoes at a time. This result is part of the folklore of tilings, and can be proved in a couple of ways, but it is by no means trivial.

We will see other examples of tile sets which have local move properties for $\mathcal{R}_{sc}$ in chapter 7 and section 8.3.

2.2.3 The set of rectangles

If we restrict our attention even further and focus only on regions which are rectangles, then there are even more sets of tiles which have a local move property. We will see examples of tile sets which have local move properties for rectangles (but not simply-connected regions in general) in chapter 4 and section 5.3.

Of course, there are trivial examples of such tile sets. Take $\mathcal{T}$ to be a set of tiles which cannot tile any rectangle. Then $\mathcal{T}$ has a local move property with respect to $\mathcal{R}_{rect}$, since for any rectangle Γ, the local move graph is empty, and thus connected. Naturally we will not be very interested in results of this form.

2.3 Proving local move properties

As we said before, local move properties are generally difficult to prove. It seems that there are basically two approaches to proving local connectivity, namely height function methods and *ad hoc* methods.

2.3.1 Height functions

Height functions are one kind of structure which can be used to prove local connectivity. Suppose Γ is a region, and τ is a tiling of this region with some set of tiles $\mathcal{T}$. A height function scheme is a procedure which assigns values (called heights) to some integer points of Γ, depending on τ. Each tiling will yield some assignment of heights (we call such a pattern of heights a *height function*). So each tiling τ has a height function f which corresponds to it, and usually this map from tilings to height functions is injective.

Generally, a height function scheme will depend upon some local condition to generate f from τ. Typically, one will begin at a point P on the boundary of Γ, and assign it some arbitrary height x. Thus $f(P) = x$ for all height functions f on Γ. Then usually there is a rule for how the height changes as one walks along the boundaries of a tile. So for instance, one could say that with each step to the east along a tile boundary, the height increases by 1, and one would have to provide similar rules for walking north, south, and west. (In this case, the rule for walking west had better be that the height decreases by 1.) In this way, each integer point which lies on the boundary of some tile in τ would receive a height. One must verify that this height function is well-defined, however. If there are several different paths from P to some other point Q in the tiling, then we must show that each of these paths yields the same height for Q. Generally all that needs to be done is to check that the height returns to its original value once we have made a loop around a single tile. If Γ is simply connected, then any loop in τ can be written as the sum of paths around individual tiles, so if the height returns to its original value on each of these smaller loops, then the same will hold for the larger loop. For this reason, height functions are most useful when trying to prove results for simply connected regions. It seems unlikely that height functions could ever be used for non-simply connected regions. One nice aspect of this approach is that the heights of points on the boundary of Γ do not depend upon τ.

To prove local connectivity using height functions, the typical strategy is usually the following. Define some notion of largeness for heights. If the heights are real numbers, this is easy, but in many cases they are elements of a group or some other structure. Then for any tiling τ, find a large height and try to show that there is a local move that can be performed there which will decrease this height. Using this idea, it may be possible to show that any height function can be reduced to a unique minimum height function. If this is true, then local connectivity holds, since every height function (or tiling) is local-move equivalent to a particular one. The trick, of course, is figuring out what height function scheme to use, and showing that there is always a local move that can decrease the height somewhere.

It seems that there are essentially two flavors of height functions, which we will call Thurston-style and Kenyon-style, in honor of their appearances in the papers [27] and [15]. Both types have some aspects in common, and the distinction between them is not always clear.

Thurston-style height functions

In [27], Thurston uses height functions to prove local connectivity for domino tilings (and also lozenge tilings on the triangular lattice). His approach involved assigning an integer to each lattice point in Γ. (The local rules were the following: Color the squares black and red in checkerboard fashion. As you move along a tile boundary, if the cell to your left is red, let the height increase by 1. If the cell to your left is black, let the height decrease by 1.) Of course, every integer point in Γ gets a height, since every point lies along a tile boundary. It turns out that the set of all height functions can be described by a set of conditions independent of tilings. Namely, a function f on the integer points of Γ is a height function if and only if the following hold:

- Values of f on the boundary of Γ obey the local condition for height functions.
- Points congruent to $(0, 0)$ mod 2 have heights congruent to 0 mod 4.
- Points congruent to $(1, 0)$ mod 2 have heights congruent to 1 mod 4.
- Points congruent to $(1, 1)$ mod 2 have heights congruent to 2 mod 4.
- Points congruent to $(0, 1)$ mod 2 have heights congruent to 3 mod 4.
- Neighboring points have values differing by at most 3.

Loosely speaking, a Thurston-style height function scheme will be one which has properties like this one. Namely, that

- Values of the height function are integers.
- Every lattice point in Γ gets a height.
- The set of height functions can be described in a way which does not have to do with tilings.

If the set of height functions can be described by a set of conditions independent of tilings, then it may be possible to define minima and maxima of height functions. If f_1 and f_2 are two height functions, then define f_{min} by the rule $f_{min}(x) =$

$\min(f_1(x), f_2(x))$. If f_{min} is a height function (and the analogous f_{max} is as well), then we can define a lattice structure on the set of tilings of Γ, by defining $f_1 \wedge f_2 = f_{min}$ and $f_1 \vee f_2 = f_{max}$. (See [26] for definitions and properties of lattices.)

We will see examples of Thurston-style height functions in chapter 4 and section 8.3.

Kenyon-style height functions

In [15], Kenyon and Kenyon use height functions to prove local connectivity for $\mathcal{R}_{sc}$ when $\mathcal{T}$ consists of a $1 \times m$ and an $n \times 1$ rectangle (thus generalizing Thurston's result). To do this, they consider the nonabelian group generated by a and b with the relations $a^m = 1$ and $b^n = 1$. Then they define a height function whose values are elements of this group. (The local rules are as follows: As you move east (resp. west) along a tile boundary, right-multiply by a (resp. a^{-1}). As you move north (resp. south) along a tile boundary, right-multiply by b (resp. b^{-1}).) Then they define the "largeness" of such a value to be its distance from $(ab)^N$ in the Cayley graph, for some sufficiently large N. (Alternatively, they could have defined the height of the original point on the boundary to be $(ab)^N$, and then declared the "largeness" of a value to be the number of multiplicative terms in its canonical expansion.)

In general, Kenyon-style height function schemes will have the following properties:

- The heights will be elements of some infinite group.
- Heights may not necessarily be defined for all lattice points of Γ.
- There is generally no description of the set of height functions which does not involve tilings.

(In the case of [15], every lattice point in Γ did get a height, but later in the paper they consider $m \times n$ and $n \times m$ rectangles, and in this case, some lattice points do not receive heights.) It seems that Kenyon-style height functions do not have some of the nice properties that Thurston-style functions have (such as the ability to take the minimum or maximum of two tilings), but they can be applied to problems where Thurston-style tilings would not work (such as cases where the tiles are large and contain lattice points in their interior).

We will see an example of a Kenyon-style height function in chapter 7.

2.3.2 Ad hoc methods

The exact definition of an *ad hoc* method is hard to say. Essentially, it is any method that does not use height functions. One such approach might be to consider the number of occurrences of the tile t_1 in a tiling. If there is always a local move (or a series of local moves) which can reduce the number of occurrences of t_1 tiles, then every tiling is local-move equivalent to a tiling which uses no t_1 tile. Then if local connectivity can be proved for tilings which involve no t_1 tile, then this proves the result. Another approach might be to consider the northernmost cell in the region. If there is a sequence of local moves which can put a particular tile there, then we may remove that portion of the region, and proceed by induction. Any technique like this, however, is easier said than done. Often such approaches (if they work at all) require careful enumeration of many different cases corresponding to all the possible arrangements of tiles near the ones of interest.

Approaches such as these seem to be most applicable when the set of regions is small (i.e., if you only want to prove local connectivity for tilings of rectangles, for instance). One notable use of ad hoc methods to prove local connectivity is a result of Donald West [30], who used local connectivity to prove the Conway-Lagarias result [7] for tiling triangles. (This problem takes place not with polyominoes, but with polyhexes, which are the analogous structure on the hexagonal lattice.)

We will see an ad hoc proof of local connectivity in section 5.3

2.4 Disproving local move properties

We know of a number of tile sets which have local move properties, but there are many more tile sets which do not have any such property. Let us see how we can succinctly disprove a local move property.

Let us consider again the case of dominoes. We mentioned earlier that the set of dominoes has a local move property for $\mathcal{R}_{sc}$. However, there is no local move property for $\mathcal{R}_{all}$, as we will show.

Let Δ_i denote the figure which is a $3 \times (2i+1)$ rectangle with the middle $1 \times (2i-1)$ rectangle removed, as shown in Figure 2-3. Observe that there are exactly two ways to tile this region (namely the way shown, and its mirror-image). Hence the local move graph has two vertices. So they must be connected by an edge. Thus, since no tile is in the same place in both tilings, converting from one tiling of Δ_i to the other must be a single local move. This must be true for all i. This requires an infinite number of local moves, which is a contradiction.

Figure 2-3: A non-simply-connected region for which local connectivity does not hold.

In general, this will be our method for disproving local move properties for any set of tiles. If we can find a region belonging to $\mathcal{R}$ which has only two tilings (and these tilings do not have tiles in common), then that region must be a local move by itself. Moreover, if we can find an infinite family of regions each of which is tileable in only two ways, then the local move property does not hold. It seems that for most tile sets, it is possible to find families of non-simply connected regions each of which has two tilings. To find simply connected regions with this property is often more difficult, and sometimes impossible.

Chapter 3

Tile invariants

3.1 Introduction

Suppose $\mathcal{T}$ is a set of tiles and Γ is a region which is tileable by these tiles. Perhaps there are many different ways to tile Γ. If so, we may still be able to say something about the types of tiles that are used. This is the idea behind tile invariants.

Let $\mathcal{T}$ be a set of n polyomino tiles, call them $t_1, \ldots, t_n$. Let $\mathcal{R}$ be a set of regions. Let Γ be a region in $\mathcal{R}$, and let τ be a tiling of this region. Define $a_i(\tau)$ to be the number of occurrences of the tile t_i in the tiling τ.

A *tile invariant* is a linear function of the $a_i(\tau)$ whose value depends only on the region $\Gamma \in \mathcal{R}$, not on the particular tiling τ. Functions whose value is invariant modulo some integer N will also be considered tile invariants. For instance, a typical tile invariant might be an equation like $a_1(\tau) + 2a_3(\tau) \equiv \text{constant} \pmod 5$, by which we mean that if τ_1 and τ_2 are two tilings of the same region Γ, then the mod 5 values of $a_1(\tau_1) + 2a_3(\tau_1)$ and $a_1(\tau_2) + 2a_3(\tau_2)$ must be equal. Of course, the "constant" in the above equation depends on the region Γ.

Following [19], we define the *tile counting group* as follows. Let $\mathbb{Z}(\mathcal{T})$ denote the group of all formal linear combinations of elements of $\mathcal{T}$. Now consider all relations of the form

$$a_1(\tau_1) \cdot t_1 + a_2(\tau_1) \cdot t_2 + \cdots + a_n(\tau_1) \cdot t_n = a_1(\tau_2) \cdot t_1 + a_2(\tau_2) \cdot t_2 + \cdots + a_n(\tau_2) \cdot t_n$$

where Γ is any region in $\mathcal{R}$, and τ_1 and τ_2 are any two tilings of that region. For instance, suppose τ_1 uses 3 copies of the tile t_1 and 2 copies of t_2, while τ_2 uses 1 copy of the tile t_1 and 4 copies of t_2. This would yield the relation

$$3t_1 + 2t_2 = t_1 + 4t_2,$$

which we may simplify to

$$2t_1 - 2t_2 = 0.$$

(Of course, this does not simplify to $t_1 = t_2$.) It should be noted that $2t_1 = 2t_2$ does not necessarily imply that there exists a region Γ which can be tiled with two copies of t_1, and also can be tiled with two copies of t_2.

Let I denote the ideal generated by the set of all such relations. Then the tile counting group is the quotient $\mathbb{Z}(\mathcal{T})/I$.

In general, the set of tile invariants (or the tile counting group) will depend upon both $\mathcal{T}$ and $\mathcal{R}$. There may be some tile invariants which hold when Γ is a rectangle, but do not hold for all simply connected regions. Similarly, there may be tile invariants which hold for all simply connected regions, but not for regions with holes. So when discussing tile invariants, we should always make sure to specify the set of regions $\mathcal{R}$.

Here is a slightly different way to think about tile invariants. Let Λ be a set of n-dimensional vectors with integer coordinates. We will say that Λ is an *integral lattice* if Λ is closed under addition and integer scalar multiplication.

Let $\mathbf{a}(\tau)$ denote the vector $(a_1(\tau), a_2(\tau), \ldots, a_n(\tau))$. Define $\Lambda_{\mathcal{T},\mathcal{R}}$ by the following rule: A vector $\mathbf{v}$ belongs to $\Lambda_{\mathcal{T},\mathcal{R}}$ if and only if there exists a region $\Gamma \in \mathcal{R}$ and two tilings τ_1, τ_2 of Γ such that $\mathbf{v} = \mathbf{a}(\tau_1) - \mathbf{a}(\tau_2)$.

Lemma 3.1 *If $\mathcal{R}$ is the set of all regions, all simply connected regions, or all rectangles, then the set $\Lambda_{\mathcal{T},\mathcal{R}}$ is an integral lattice.*

Proof: First observe that if $\mathbf{v} \in \Lambda_{\mathcal{T},\mathcal{R}}$, then we also have $-\mathbf{v} \in \Lambda_{\mathcal{T},\mathcal{R}}$ just by switching the roles of τ_1 and τ_2 in the definition. All that remains to be shown is that $\Lambda_{\mathcal{T},\mathcal{R}}$ is closed under addition. Let $\mathbf{v}$ and $\mathbf{v}'$ belong to $\Lambda_{\mathcal{T},\mathcal{R}}$. Let $\mathbf{v} = \mathbf{a}(\tau_1) - \mathbf{a}(\tau_2)$ and let

$\mathbf{v}' = \mathbf{a}(\tau_1') - \mathbf{a}(\tau_2')$, where τ_1 and τ_2 are tilings of Γ, and τ_1' and τ_2' are tilings of Γ'.

The next step is to paste the regions Γ and Γ' together in some way to create a new region Γ^*. The only complication is that we want the new region to belong to $\mathcal{R}$. If $\mathcal{R}$ is the set of all regions, then there is no issue. If $\mathcal{R}$ is the set of simply connected regions, then let γ be the right-most cell in the bottom row of Γ, and let γ' be the left-most cell in the top row of Γ'. Now affix the bottom of γ to the top of γ'. The resulting region is then simply connected (see Figure 3-1). (We'll deal with rectangles later.)

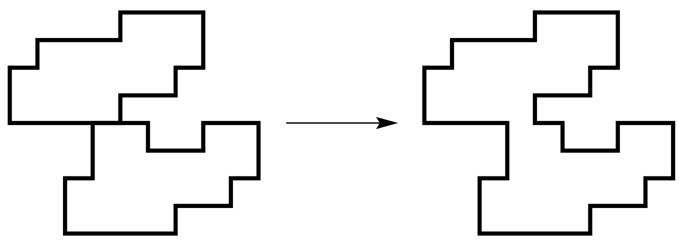


Figure 3-1: Pasting together copies of Γ and Γ' to form the simply connected region Γ^*.

For $i = 1, 2$, let τ_i^* denote the tiling of Γ^* obtained by pasting together τ_i and τ_i'. Then $\mathbf{a}(\tau_i^*) = \mathbf{a}(\tau_i) + \mathbf{a}(\tau_i')$. It follows that $\mathbf{v} + \mathbf{v}' = \mathbf{a}(\tau_1^*) - \mathbf{a}(\tau_2^*)$, hence $\mathbf{v} + \mathbf{v}' \in \Lambda_{\mathcal{T},\mathcal{R}}$, as desired.

Now, if $\mathcal{R}$ is the set of rectangles, we have to be a little bit more clever. Suppose Γ is a $p \times q$ rectangle, and Γ' is a $p' \times q'$ rectangle. Stack p' copies of Γ on top of one another, forming a $pp' \times q$ rectangle, and place these next to a stack of p copies of Γ'. This forms a $pp' \times (q + q')$ rectangle Γ^*. Now let τ_1^* denote the tiling of Γ^* where every subrectangle is tiled like τ_1 or τ_1'. Let τ_2^* denote the tiling of Γ^* where *one* subrectangle is tiled like τ_2 and *one* subrectangle is tiled like τ_2', and all the rest are tiled like τ_1 or τ_1'. Then $\mathbf{v} + \mathbf{v}' = \mathbf{a}(\tau_1^*) - \mathbf{a}(\tau_2^*)$ as before. ■

Under this interpretation, a tile invariant is a vector $\mathbf{w}$ such that the dot product $\mathbf{w} \cdot \mathbf{v}$ equals 0 for all $\mathbf{v} \in \Lambda_{\mathcal{T},\mathcal{R}}$ (or the dot product always equals 0 mod N, for some N). Of course there are some trivial examples of $\mathbf{w}$ for which this holds; for instance, take $\mathbf{w} = \mathbf{0}$. Naturally, we will not be particularly interested in tile invariants of this type.

The tile counting group in this case is the quotient $\mathbb{Z}^n/\Lambda_{\mathcal{T},\mathcal{R}}$.

Determining the tile invariants is equivalent to determining the integral lattice $\Lambda_{\mathcal{T},\mathcal{R}}$. Generally we will do this by constructing a basis for $\mathbb{Z}^n$ for which $\Lambda_{\mathcal{T},\mathcal{R}}$ is particularly nice.

3.2 Proving tile invariants

It seems that there are essentially three standard techniques for proving tile invariants: coloring arguments, boundary-word arguments, and local connectivity. The first two of these are very similar to techniques for proving untileability of certain regions. Let us consider each method briefly.

3.2.1 Coloring arguments

The name "coloring argument" comes from the classic problem of tiling a region with dominoes. Color the region with two colors, as you would color a checkerboard. Since each domino covers one red square and one black square, it follows that any tileable region must contain an equal number of red and black squares. If a region has an unequal number of red and black squares, then we say that a coloring argument rejects tileability. In general, a coloring argument is one in which values are assigned to the cells of the grid, and the total value covered by each tile is considered. For the domino problem, we would assign the value of 1 to each red cell and assign -1 to each black cell, and observe that each domino covers cells whose total value is 0. Assigning numerical values instead of colors allows for greater flexibility, but for historical reasons we still use the term "coloring argument". Frequently these values will be numbers modulo some integer N.

In the context of tile invariants, our goal would be to find a numbering scheme such that the value covered by a tile depends only on the tile's shape, and not its location. For instance, we might have a scheme in which every possible t_1 tile covers cells whose sum is 1 (mod 4), and every possible t_2 tile covers cells whose sum is -1 (mod 4). Assuming there are no other tiles in $\mathcal{T}$, then the sum of all values in Γ is equivalent to $1 \cdot a_1(\tau) + (-1) \cdot a_2(\tau) \pmod 4$. Since the sum of all values in Γ is a constant (it does not depend on τ), this would then be a tile invariant.

Notice that such a technique proves tile invariants for $\mathcal{R}_{all}$. This is both good and bad. It is good because it can prove tile invariants for regions which contain holes. It is bad because it means that this technique cannot be used to prove tile invariants for $\mathcal{R}_{sc}$ that do not hold for $\mathcal{R}_{all}$.

3.2.2 Boundary word arguments

The technique of boundary words was invented by Conway and Lagarias in their groundbreaking paper [7]. In that paper, the goal was to prove untileability of certain regions. The idea, roughly speaking, is to assign letters to the (directed) edges of the grid, and consider the words traced out as one walks along the boundary of a simply connected region. In [7], they label horizontal edges a, and direct them to the east, and label vertical edges b, and direct them to the north[1]. Then to form the boundary word of a region Γ, begin at an arbitrary point on the boundary of Γ, and proceed counterclockwise. With every edge traversed in the direction of its arrow, write down the appropriate letter. With every edge traversed in the opposite direction, write down the letter inverse. Treat the resulting word $w(\Gamma)$ as an element of the free (nonabelian) group on generators a and b.

Each tile t_i corresponds to a word $w(t_i)$ in this scheme. The key observation is that if Γ is a tileable region, then $w(\Gamma)$ can be expressed as a product of conjugates of the words $w(t_i)$. (See [7] for a full explanation of this.) Let H denote the group generated by the words $w(t_i)$ and their conjugates. Notice that every element of H corresponds to a closed loop (i.e., it ends where it begins). Also notice that $w(\Gamma)$ belongs to H for every simply connected tileable region Γ. Let C denote the group of *all* words which correspond to closed loops. Then define the *tile homotopy group* to be the quotient C/H. If Γ is a region, and $w(\Gamma)$ is not equal to the identity in the group C/H, then this proves that Γ is not tileable. (Of course, proving that an element of C/H is not the identity usually requires some clever trick (such as considering an

[1] Actually they use the letters A and U, which stand for "across" and "up".

appropriate quotient of this group), so this method is not completely automatic.)

Inspired by [23], we will use this method in a slightly different way so that it applies to tile invariants.

As before, we will assign labels and directions to the edges of the unit grid so that each simply connected region Γ has an associated boundary word $w(\Gamma)$. We may assign these labels in a non-traditional way, but for now the only important thing is for the boundary word of a tileable region to be expressible as a product of conjugates of boundary words of tiles. Now we will take these boundary words $w(\Gamma)$ and use them to draw closed paths $\pi(\Gamma)$ in an alternate universe. Naturally, we will want to interpret the letters in the word differently than we did before, or else we will not have accomplished anything. In the context of [23], we imagine that there are two pedestrians walking in two different cities, talking to each other on their cell phones. The first pedestrian walks around some region Γ in his city, and tells his friend the labels of the streets and avenues he walks along. These labels provide instructions for where the second pedestrian should go in her city. We let $\pi(\Gamma)$ denote the path traced out by the second pedestrian.

At this point, what we need to verify is that every possible tile placement t_i in the original universe generates a word $w(t_i)$ which becomes a closed path $\pi(t_i)$ in the alternate universe. It follows that if Γ is a tileable region, then $\pi(\Gamma)$ will be a closed path in the alternate universe. In some sense, if Γ is tiled by tiles t_i, then the alternate path $\pi(\Gamma)$ can be built up from conjugates of the $\pi(t_i)$.

At this point, we employ coloring arguments in the alternate universe. Typically this involves taking a weighted signed area. If π is a closed path, and c is a cell, then the winding number $\omega_\pi(c)$ denotes the number of times π wraps around the cell counterclockwise, minus the number of times it wraps around clockwise. (See any book on complex analysis for a better treatment of winding numbers.) Clearly, if π is a finite closed path, then all but finitely many of these winding numbers will be 0. Say we assign a value $\iota(c)$ to each cell of the grid. Then the weighted signed area corresponding to π is defined as $\sum_c \iota(c)\omega_\pi(c)$, where the sum is taken over all cells in the grid. Observe that the weighted signed area is additive—if a closed path π can be written as the sum of closed paths π_1 and π_2, then the weighted signed area of π is the sum of the weighted signed areas of π_1 and π_2.

Thinking about tilings again, suppose Γ is a region in the original universe which is tiled by tiles t_i. Then $\pi(\Gamma)$ can be built up from the paths $\pi(t_i)$ corresponding to its constituent tiles. Hence the weighted signed area of $\pi(\Gamma)$ will be the sum of the weighted signed areas of its constituent $\pi(t_i)$'s. So, if our weighting scheme always assigns the same weighted signed area ρ_i to every possible placement of a t_i tile, then this may allow us to prove a tile invariant.

Notice that this method depends upon the region Γ being simply connected. As such, it seems like this technique can only work to prove tile invariants which hold for $\mathcal{R}_{sc}$. We will see examples of this technique in section 8.2 and chapters 9 and 10.

3.2.3 Local connectivity

The third way to prove tile invariants is by using local connectivity. If there is a local move property for $\mathcal{T}$ and $\mathcal{R}$, then there is some set of local moves $\mathcal{L}$ so that one can convert any tiling to any other by means of these moves. It just remains to show that the desired tile invariants hold for each of the local moves.

This technique is very easy to use, assuming local connectivity has already been established. Naturally, the hard part is proving local connectivity.

3.3 Dominoes

Let us consider a specific example. Let $\mathcal{T}$ be the set of dominoes. Let t_1 be the horizontal domino, and let t_2 be the vertical domino. The tile invariants for dominoes have been known for a long time, and are part of the folklore of domino tilings, even if the notion of a tile invariant is a relatively new one.

Figure 3-2: Dominoes.

Let us consider the set of all regions. Fix a region $\Gamma \in \mathcal{R}_{all}$. Suppose it has area c. Then Γ must be tiled by $c/2$ tiles, if it can be tiled at all. Hence $a_1(\tau) + a_2(\tau)$ is a tile invariant, because its value is always $c/2$, regardless of the tiling τ.

There is also another tile invariant, namely that $a_1(\tau)$ is invariant mod 2. Consider the following coloring argument. Assign the value 0 to those grid cells with an even x-coordinate, and assign the value 1 to those grid cells with an odd x-coordinate.

$$\begin{array}{ccccccc} 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{array}$$

Let d be the sum of the values of all the cells in Γ, taken mod 2. Notice that a vertical domino always covers two cells whose sum is 0 mod 2, while a horizontal domino always covers two cells whose sum is 1 mod 2. Hence we have that $a_1(\tau) \equiv d$ mod 2, thus $a_1(\tau)$ is a mod-2 tile invariant.

Let us now devise a new basis for $\mathbb{Z}^n$ which will allow us to show that there are no other tile invariants which do not follow from these. Define

$$b_1(\tau) = a_1(\tau) + a_2(\tau)$$
$$b_2(\tau) = a_1(\tau)$$

Notice that $b_1(\tau)$ and $b_2(\tau)$ have the same integer span as $a_1(\tau)$ and $a_2(\tau)$. In other words, the matrix which transforms an a-vector into a b-vector, and the inverse of this matrix, both have integer entries. We have already shown that $b_1(\tau)$ is invariant, and that $b_2(\tau)$ is invariant mod 2. Hence any element of the integral lattice $\Lambda_{\mathcal{T},\mathcal{R}}$ must have the form $(0, 2c)$ in the b-basis. In order to show that these are the best possible invariants, it suffices to construct an example of a region Γ with tilings τ_1 and τ_2 such that $b_2(\tau_1)$ and $b_2(\tau_2)$ differ by exactly 2. This will show that the b-vector $(0, 2)$ belongs to $\Lambda_{\mathcal{T},\mathcal{R}}$, and thus we will have determined the integral lattice completely.

In this instance, we can take Γ to be a 2×2 square. It can be tiled with two vertical dominoes ($b_2(\tau) = 0$) or with two horizontal dominoes ($b_2(\tau) = 2$).

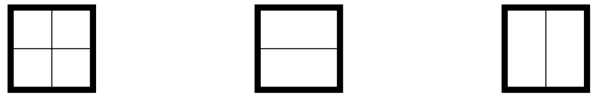


Figure 3-3: A region with two tilings which demonstrates that b_2 may vary by 2.

In this case, the region which generated the b-vector $(0, 2)$ was a rectangle. Hence $\Lambda_{\mathcal{T},\mathcal{R}}$ is the same regardless of whether $\mathcal{R}$ is the set of all regions, all simply-connected regions, or all rectangles.

Since we have b_1 invariant mod ∞, and b_2 invariant mod 2, it follows that the tile counting group in this case will be isomorphic to $\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Now let us proceed to a more difficult example.

3.4 T-tetrominoes

Consider the case of T-tetrominoes. Let t_1 be the upward-pointing tile, let t_2 be the rightward-pointing tile, let t_3 be the downward-pointing tile, and let t_4 be the leftward-pointing tile.

Figure 3-4: T-tetrominoes.

In this case, we will present our new basis first, and then present the tile invariants within that context. Let

$$
\begin{aligned}
b_1(\tau) &= a_1(\tau) + a_2(\tau) + a_3(\tau) + a_4(\tau) \\
b_2(\tau) &= a_1(\tau) + a_2(\tau) \\
b_3(\tau) &= a_1(\tau) + a_4(\tau) \\
b_4(\tau) &= a_1(\tau).
\end{aligned}
$$

Again notice that this transformation, and its inverse, are integer transformations.

Theorem 3.2 *We have that $b_1(\tau)$ is invariant (mod ∞), $b_2(\tau)$ is invariant mod 4, and $b_3(\tau)$ is invariant mod 4.*

Proof: The first assertion is easy; $b_1(\tau)$ always equals the area of Γ divided by 4. As for the second assertion, number the cells of the grid according to the following pattern (all taken mod 32).

7	3	15	11	23	19	31	27	7	3	15
27	7	3	15	11	23	19	31	27	7	3
31	27	7	3	15	11	23	19	31	27	7
19	31	27	7	3	15	11	23	19	31	27
23	19	31	27	7	3	15	11	23	19	31

One can verify that any placement of a tile of type t_1 or t_2 will cover cells summing to 8 mod 32, while any tile of type t_3 or t_4 will cover cells summing to 0 mod 32. Hence the sum of all the cells in Γ will equal $8 \cdot b_2(\tau)$ (mod 32), thus $b_2(\tau)$ is invariant modulo 4. The argument for $b_3(\tau)$ is essentially the same, just flipped horizontally. ∎

These invariants hold for all regions.

Theorem 3.3 *The invariants in Theorem 3.2 determine the integral lattice $\Lambda_{\mathcal{T},\mathcal{R}}$ completely for $\mathcal{R}_{all}$ and $\mathcal{R}_{sc}$.*

Proof: First we will give an example which shows that b_4 can vary freely.

Consider the following region Γ (Figure 3-5), and the two tilings shown. The first tiling contains one copy of t_1 and one copy of t_3, which in the b-basis is $(b_1, b_2, b_3, b_4) = (2, 1, 1, 1)$. The second tiling contains one copy of t_2 and one copy of t_4, which in the b-basis is $(2, 1, 1, 0)$. Thus the b-vector $(0, 0, 0, 1)$ belongs to $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$. So b_4 can vary freely.

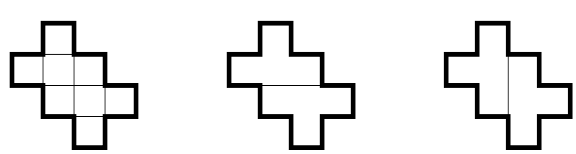


Figure 3-5: A region with two tilings which demonstrates that b_4 can vary freely.

Now we will give an example which shows that b_2 can vary by exactly 4.

Consider the region Γ in Figure 3-6, and the two tilings shown. The first tiling is $(5,3,2,2)$ in the a-basis, and in the b-basis it is $(12,8,7,5)$. The second tiling is $(2,2,3,5)$ in the a-basis, and in the b-basis it is $(12,4,7,2)$. So the b-vector $(0,4,0,3)$ belongs to $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$. We already showed that $(0,0,0,1)$ belongs to $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$, and since $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$ is an integral lattice, we must have that $(0,4,0,0)$ belongs to $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$. Thus b_2 can vary by exactly 4.

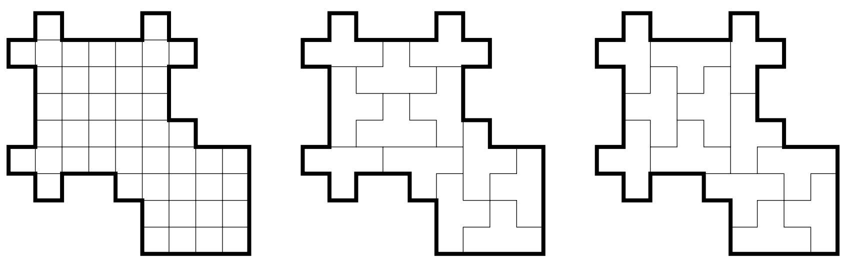


Figure 3-6: A region with two tilings which demonstrates that b_2 can vary by exactly 4.

This same argument, flipped horizontally, demonstrates that b_3 can vary by exactly 4. Since $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$ contains the b-vectors $(0,4,0,0)$, $(0,0,4,0)$, and $(0,0,0,1)$, it must contain all vectors of the form $(0,4c_1,4c_2,c_3)$. Theorem 3.2 states that only vectors of that form are allowed, hence we have found $\Lambda_{\mathcal{T},\mathcal{R}_{sc}}$ exactly. ■

So the tile counting group for $\mathcal{R}_{sc}$ and $\mathcal{R}_{all}$ is isomorphic to $\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$.

If we could have found rectangular regions which yielded the necessary b-vectors, then we would have proved that these tile invariants were the best possible for $\mathcal{R}_{rect}$. However, we will see in chapter 4 that stronger tile invariants hold when we limit ourselves to tilings of rectangles.

Chapter 4

Tiling with T-tetrominoes

4.1 Introduction

Recall that a *T-tetromino* is the figure formed by four unit squares arranged as shown in Figure 4-1. We allow all four orientations.

Figure 4-1: T-tetrominoes.

For the first part of this chapter, we will be looking only at tilings of rectangles by T-tetrominoes. We will make some observations about the structure of such tilings, with the goal being to prove a local-connectivity result. Later, we will extend these results to a somewhat more general class of regions, and we will prove a result concerning the number of tilings of such regions.

Two natural local moves for T-tetrominoes are shown in Figure 4-2. We call them the 2-move and the 4-move. One of our main results is the following.

Figure 4-2: Local 2-move and local 4-move.

Theorem 4.1 *The set of T-tetrominoes has a local-move property for $\mathcal{R}_{rect}$. Specifically, every rectangle Γ has local connectivity with respect to the 2-move and 4-move.*

This result was conjectured in [18] to hold for all simply connected regions. Later on, in section 4.8, we extend this theorem to a more general class of regions and show that the conjecture does not hold in full generality.

4.2 Tiling rectangles with T-tetrominoes

Without loss of generality, let Γ be a rectangle which is situated in the first quadrant of the Cartesian plane, with one corner at $(0, 0)$. Let a *type-A* point be a point whose coordinates are congruent mod 4 to (0,0) or (2,2), and let a *type-B* point be a point whose coordinates are congruent mod 4 to (0,2) or (2,0). A segment of length 1 is called a *cut* if there is no valid tiling of Γ in which a tile crosses that segment. A point is called *cornerless* if there is no valid tiling of Γ in which that point is one of the eight corners of a tile.

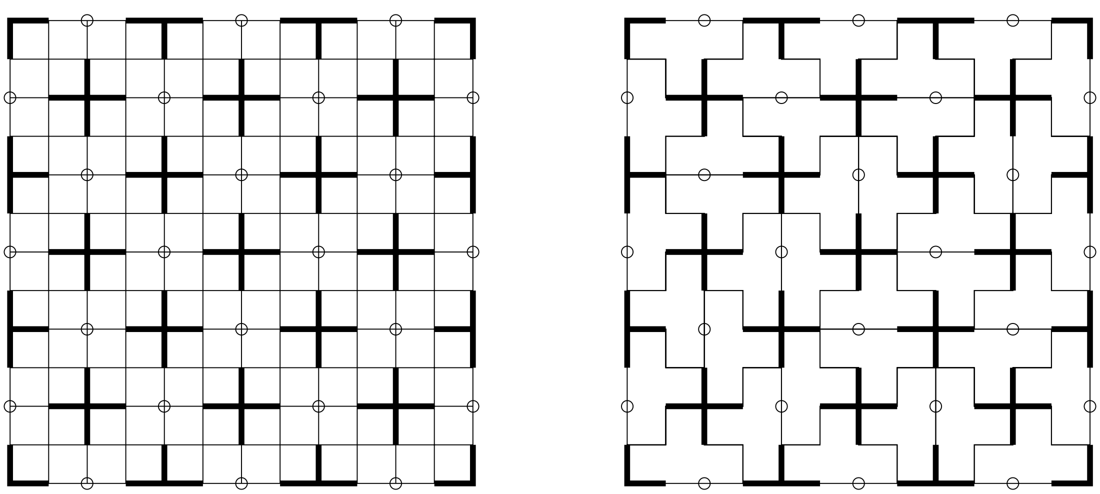


Figure 4-3: The dark lines are cuts. Circles are cornerless points.

In [28], Walkup proves the following property of T-tetromino tilings of rectangles (see Figure 4-3).

Theorem 4.2 (Walkup) *If an $m \times n$ rectangle can be tiled by T-tetrominoes, then both m and n must be divisible by 4. Furthermore, all segments incident to type-A points are cuts, and all type-B points are cornerless.*

From now on, we will only be concerned with rectangles having sides divisible by 4, since all other rectangles are untileable.

Define a *block* to be a 2×2 square whose corners have even coordinates. The following lemma is immediate by inspection from the structure of cuts and cornerless points.

Lemma 4.3 *In any tiling of a rectangle by T-tetrominoes, each tile contains three squares from one block and one square from an adjacent block. Similarly, each block contains three squares from one tile and one square from another tile.*

4.3 Chain graphs

Define an *antiblock* to be a 2×2 square whose corners have odd coordinates. Color the antiblocks white and gray in checkerboard fashion, so that antiblocks centered at type-A points are gray and those centered at type-B points are white.

For a $4m \times 4n$ rectangle Γ, let V_Γ be the set of points in Γ which have odd coordinates. Say that a directed graph on the vertices V_Γ is a *chain graph* if it satisfies the following properties:

- every edge connects vertices that are two units apart (either vertically or horizontally),
- every vertex has indegree 1 and outdegree 1, and
- every white antiblock contained in Γ borders exactly two edges of the graph, and these edges are non-adjacent.

Let $\mathcal{C}_\Gamma$ denote the set of all chain graphs of a region Γ. Let $\mathcal{Y}_\Gamma$ denote the set of all tilings of Γ.

Theorem 4.4 *For any $4m \times 4n$ rectangle Γ, we have $|\mathcal{C}_\Gamma| = |\mathcal{Y}_\Gamma|$.*

The proof is based on an explicit bijection $\phi : \mathcal{Y}_\Gamma \to \mathcal{C}_\Gamma$ defined as follows.

Let $\tau \in \mathcal{Y}_\Gamma$ be a tiling. Notice that each vertex in V_Γ lies in the middle of some block. By Lemma 4.3, each tile in τ contains three squares from one block and one square from an adjacent block. Call these blocks the *primary* and *secondary* blocks of the tile respectively. For each tile, draw a directed edge from its primary block to its secondary block, and define $\phi(\tau)$ to be the directed graph which results (see Figure 4-4).

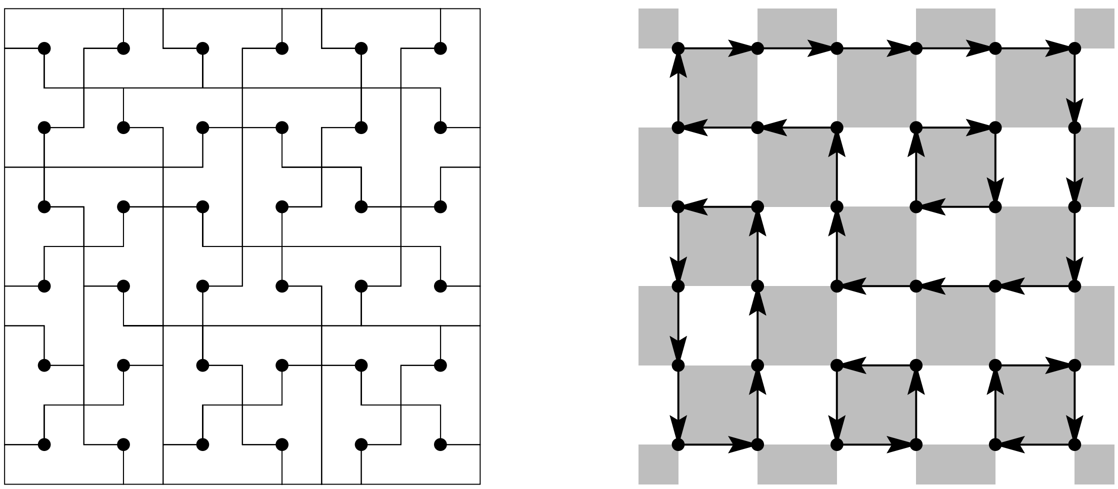


Figure 4-4: A tiling τ, and the chain graph $\phi(\tau)$.

Theorem 4.4 follows immediately from the following lemma.

Lemma 4.5 *For any $4m \times 4n$ rectangle Γ, the map ϕ defined above is a bijection between $\mathcal{Y}_\Gamma$ and $\mathcal{C}_\Gamma$.*

Proof: First let us show that $\phi(\tau)$ is a chain graph. It is clear from the definition, and from Lemma 4.3, that every vertex will have indegree 1 and outdegree 1, and that edges will only connect vertices which are two units apart. As for the third restriction, consider a type-B point not on the boundary. Up to rotations and reflections, the tiles surrounding it must look like one of the two possibilities shown in Figure 4-5. Thus there will be exactly two edges bordering the associated white antiblock, and they will be non-adjacent. Hence $\phi(\tau)$ is a chain graph for all τ.

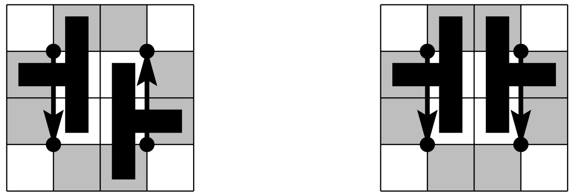


Figure 4-5: The two possibilities for a type-B point.

Notice that each tile corresponds to an edge in this graph. For each edge, there is only one possible tile placement which yields that edge and is consistent with the cuts and cornerless points. Hence the map ϕ is injective.

What remains to be shown is that every chain graph is equal to $\phi(\tau)$ for some tiling τ. As we just observed, for each edge there is only one possible tile placement that can yield that edge. So any chain graph will yield a collection of tile placements. It remains to be checked that these tiles cover all of Γ and do not overlap. Since each vertex has outdegree 1, the number of edges equals the number of blocks, so the total area of the tiles will equal the area of Γ. Thus it will be sufficient to verify that the tiles do not overlap.

Assume there are two tiles which overlap. Let us assume the overlap occurs in the block containing the squares A, B, D, and E (see Figure 4-6). Without loss of generality, we may take one of the tiles to be the one covering squares B, D, E, and F. Since each vertex has indegree 1 and outdegree 1, the tile which overlaps this one must contain only one square from this block, hence the overlap must occur at E. There are two possible tiles which cover E. First there is the tile which covers C, E, F, and G. If we have this, then the graph must contain both an edge and its opposite. This violates the rule about what a white antiblock may border. The other possibility is the tile which covers E, H, I, and J. In this case, the graph must contain two adjacent edges both on the same white antiblock, which again violates the constraint. Thus there can be no overlaps, which proves that every chain graph is $\phi(\tau)$ for some tiling τ. ∎

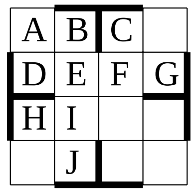


Figure 4-6: How an overlap may occur.

4.4 Height functions

Let us call a point having coordinates congruent mod 4 to (0,0) a *type-A0* point. Similarly, a point congruent to (2,2) will be called a *type-A1* point. (Points congruent to either (0,2) or (2,0) will still be called *type-B* points.)

For a $4m \times 4n$ rectangle Γ, let W_Γ be the set of points in Γ which have even coordinates. Let $\partial\Gamma$ denote the set of boundary points of Γ. Say that a function $f : W_\Gamma \to \mathbb{Z}$ is a *height function* if it satisfies the following properties:

- $f(x) = 0$ for all $x \in \partial\Gamma$,
- $f(x)$ is an even integer for all type-A0 points x,
- $f(x)$ is an odd integer for all type-A1 points x, and
- $|f(x) - f(y)| \leq 1$ whenever x and y are adjacent (at a distance of two units).

Let $\mathcal{H}_\Gamma$ denote the set of all height functions of a region Γ.

Theorem 4.6 *For any $4m \times 4n$ rectangle Γ, we have $|\mathcal{H}_\Gamma| = |\mathcal{Y}_\Gamma|$.*

We define a map $\psi : \mathcal{C}_\Gamma \to \mathcal{H}_\Gamma$ as follows. Let $C \in \mathcal{C}_\Gamma$ be a chain graph. Define a function f° on the faces of C by the following rules. Let f° have the value 0 on the unbounded face of C. As we pass an edge of the graph, if the edge points to the right, let the value of f° increase by 1. (Similarly, if the edge points to the left, let the value of f° decrease by 1.) Now define $f : W_\Gamma \to \mathbb{Z}$ by letting $f(x)$ equal the value of f° on the face in which x lies (see Figure 4-7). Define $\psi(C)$ to be this function f.

Theorem 4.6 follows immediately from Theorem 4.4 and the following lemma.

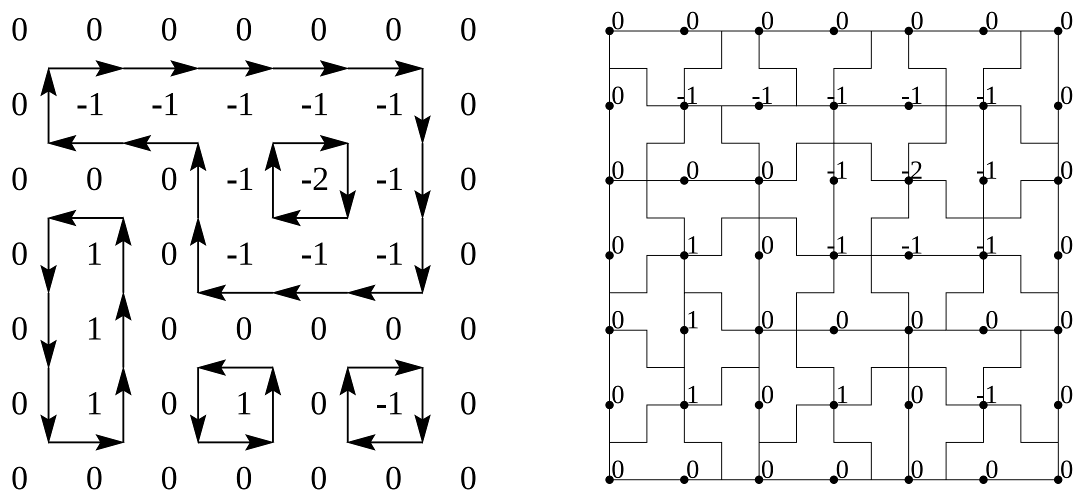


Figure 4-7: A chain graph C, and the function $f = \psi(C)$.

Lemma 4.7 *For any $4m \times 4n$ rectangle Γ, the map ψ defined above is a bijection between $\mathcal{C}_\Gamma$ and $\mathcal{H}_\Gamma$.*

Proof: Let C be a chain graph, and let f be $\psi(C)$ as defined above. Let us first show that the function f is well-defined. If it is not, then there must exist some closed path through the faces of the graph such that the net change in the value of f° is non-zero. This means that upon going around this path counterclockwise, we cross more right-pointing edges than left-pointing edges, say. Therefore more edges leave the area enclosed by the path than enter that area. But this is impossible since every vertex has equal indegree and outdegree, so the net flow out of any region must be zero. Hence f is a well-defined function on W_Γ.

Next, let us verify that f is a valid height function. Points $x \in \partial\Gamma$ lie in the unbounded face of C, hence $f(x) = 0$ for such points. And if x and y are adjacent points, then they lie either in the same face of C or in adjacent faces of C, hence the difference between $f(x)$ and $f(y)$ is at most 1. Now let us verify the other two statements. As one travels from a type-A0 point x to another type-A0 point y which is 4 units away, one passes through the middle of a white antiblock (see Figure 4-8). In doing so, one crosses either 0 or 2 edges of C, hence the value of f° will have changed twice, or not at all, so $f(x)$ and $f(y)$ will have the same parity. Since (0,0) is a type-A0 point, and $f((0,0)) = 0$, it follows that $f(x)$ will be even for all type-A0 points x. By the same argument, all type-A1 points must have the same parity as each other. And (2,2) is a type-A1 point with $f((2,2)) = \pm 1$, so $f(x)$ will be odd for all type-A1 points x. Thus f is in fact a height function.

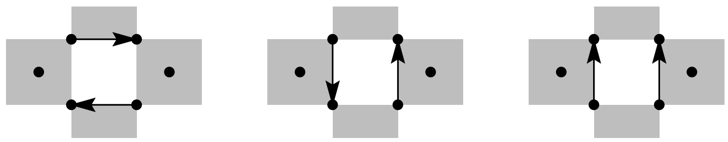


Figure 4-8: Two type-A0 points, and what might lie between them.

Given a height function $f = \psi(C)$, one can uniquely reconstruct the chain graph C by inserting directed edges in the places where the value of f increases or decreases. Hence ψ is an injective map. It remains to be shown that every height function f is equal to $\psi(C)$ for some valid chain graph C.

Take a height function f, and insert directed edges along the boundaries where the value of f increases or decreases. Call this graph C. Consider a vertex of C. To one corner of it, there is a type-A0 point x_0, on the opposite corner is a type-A1 point x_1, and the remaining two corners are type-B points y_0 and y_1. Since $f(x_0)$ is even, and $f(x_1)$ is odd, these values must differ by exactly 1. Without loss of generality, assume $f(x_0) = h$ and $f(x_1) = h + 1$. Then both $f(y_0)$ and $f(y_1)$ must be h or $h + 1$ as well. Up to rotations, the situation must look like one of the possibilities in Figure 4-9. Thus the vertex in question will have indegree 1 and outdegree 1.

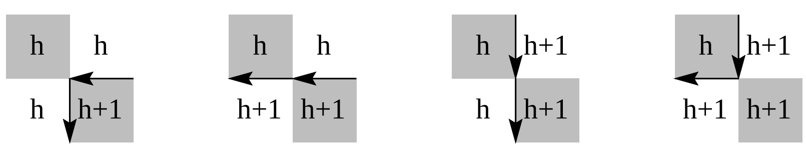


Figure 4-9: The possibilities for a vertex of C.

Now consider a type-B point y, which corresponds to a white antiblock. Let $f(y) = h$, and assume without loss of generality that h is even. If z_1 and z_2 are the two type-A0 points adjacent to y, then we must have $f(z_1) = f(z_2) = h$. If z_3 and z_4 are the two type-A1 points adjacent to y, then we must have $f(z_3) = h \pm 1$ and $f(z_4) = h \pm 1$, not necessarily the same (see Figure 4-10). So this white antiblock will border exactly two non-adjacent edges of C.

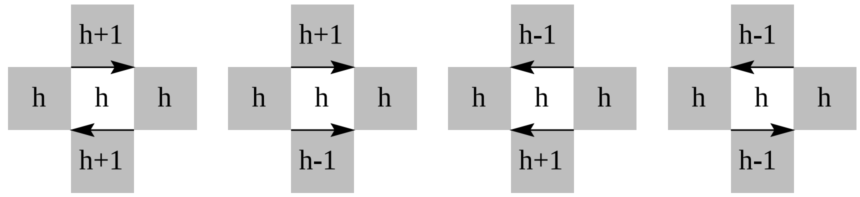


Figure 4-10: The possibilities for a white antiblock.

Hence the graph C constructed in this way from a height function f is indeed a chain graph, and $\psi(C) = f$. This completes the proof. ■

For ease of notation, define $\zeta(\tau) = \psi(\phi(\tau))$. For a $4m \times 4n$ rectangle Γ, the map ζ is the canonical bijection between $\mathcal{Y}_\Gamma$ and $\mathcal{H}_\Gamma$.

Lemma 4.8 *Let Γ be a $4m \times 4n$ rectangle and let $\tau_1, \tau_2 \in \mathcal{Y}_\Gamma$ be tilings of Γ. The tilings τ_1 and τ_2 differ by a 2-move if and only if the height functions $\zeta(\tau_1)$ and $\zeta(\tau_2)$ differ by 1 on some type-B point, and are the same everywhere else. The tilings τ_1 and τ_2 differ by a 4-move if and only if the height functions $\zeta(\tau_1)$ and $\zeta(\tau_2)$ differ by 2 on some type-A point, and are the same everywhere else.*

Proof: By inspection of the structure of cuts and cornerless points, one sees that the 2-move must be centered at a type-B point, and the 4-move must be centered at a type-A point. From Figure 4-11, one can see that if τ_1 and τ_2 differ by a 2-move, then the height functions $\zeta(\tau_1)$ and $\zeta(\tau_2)$ differ by 1 in their values on the corresponding type-B point. Similarly, if τ_1 and τ_2 differ by a 4-move, then the height functions $\zeta(\tau_1)$ and $\zeta(\tau_2)$ differ by 2 in their values on the corresponding type-A point.

As for the converse, suppose there are height functions f_1 and f_2 which are identical everywhere, except $f_1(y) = h$ and $f_2(y) = h + 1$ for some type-B point y. Thus the value of f_1 (or f_2) on the neighbors of y must be $h, h+1, h$, and $h+1$ (since they must alternate even and odd). Hence the picture must look like the bottom left of Figure 4-11, possibly rotated. Going backwards, we see what the chain graph and the tiling must then look like, and that in fact, $\zeta^{-1}(f_1)$ and $\zeta^{-1}(f_2)$ differ by a 2-move.

Similarly, suppose there are height functions f_1 and f_2 which are identical everywhere, except $f_1(x) = h + 1$ and $f_2(x) = h - 1$ for some type-A point x. Thus

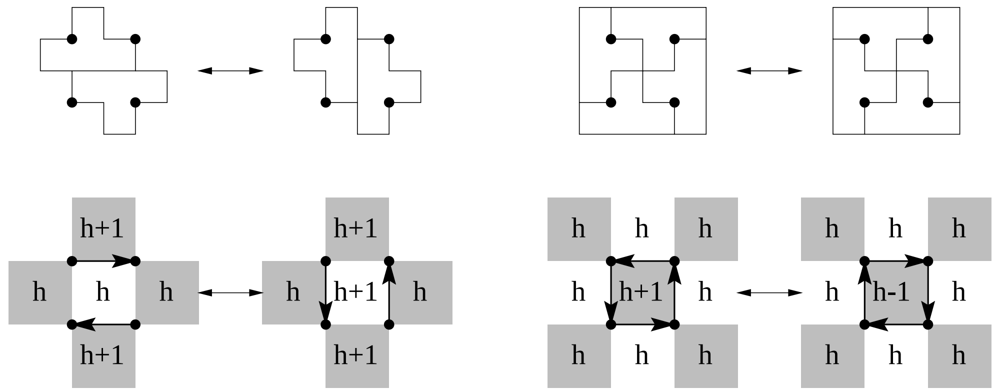


Figure 4-11: The 2-move and 4-move, and their effect on $\zeta(\tau)$.

$f_1(y) = f_2(y) = h$ for all neighbors y of x. Hence the picture must look like the bottom right of Figure 4-11. Going backwards, we see what the chain graph and the tiling must then look like, and that in fact, $\zeta^{-1}(f_1)$ and $\zeta^{-1}(f_2)$ differ by a 4-move. ■

For height functions $f_1, f_2 \in \mathcal{H}_\Gamma$, say that f_1 and f_2 differ by a 2-move (or 4-move) if the tilings $\zeta^{-1}(f_1)$ and $\zeta^{-1}(f_2)$ differ by a 2-move (or 4-move). By the previous lemma, one can see that performing a 2-move on a height function f is equivalent to increasing or decreasing its value by 1 at some type-B point. Similarly, performing a 4-move is equivalent to increasing or decreasing the value of f by 2 at some type-A point. Of course, such moves may only be applied if the function that results is a valid height function.

4.5 Local connectivity from height functions

Theorem 4.1 will easily follow from the following lemma.

Lemma 4.9 *Let Γ be a $4m \times 4n$ rectangle, and let $f_1, f_2 \in \mathcal{H}_\Gamma$ be height functions. It is always possible to convert f_1 into f_2 by performing a sequence of 2-moves and 4-moves.*

Proof: For a $4m \times 4n$ rectangle Γ, let f_0 be the height function which is 1 on the type-A1 points of Γ, and 0 everywhere else. We would like to show that every height function f can be transformed into f_0. If every height function can be transformed into f_0, it follows that any height function can be transformed into any other. Suppose $f(x) > 1$ for some x. Let h be the largest value that f attains. Suppose there is a type-B point y which attains this value. Then f must take the values $h, h-1, h$, and $h-1$ on the neighbors of y. So we can perform a 2-move to change $f(y)$ to $h-1$ and still have a valid height function. We do this for all type-B points at which f attains the value h. Now look at any remaining (type A0 or A1) point x having $f(x) = h$. We must have $f(z) = h-1$ for the neighbors z of x, since there are no type-B points remaining for which $f(z) = h$. So we can perform a 4-move to change $f(x)$ to $h-2$. We do this for every point where f attains the value h. Now the largest value which appears is at most $h-1$, and we repeat the procedure until we have $f(x) \leq 1$ for all x.

We do a similar thing for points where $f(x) < 0$, increasing them until $f(x) \geq 0$ for all x. At this point, all points will have the value 0 or 1 (in particular, $f(x) = 0$ for all type-A0 points x, and $f(x) = 1$ for all type-A1 points). It just remains to set $f(y) = 0$ for all type-B points y, which can be done by a sequence of 2-moves. This finishes the procedure, proving the lemma. ∎

4.6 The lattice structure on height functions

There is a natural partial order on $\mathcal{H}_\Gamma$. If $f_1, f_2 \in \mathcal{H}_\Gamma$ are height functions, we say $f_1 \leq f_2$ iff $f_1(x) \leq f_2(x)$ for all points x. This partial order can be extended to tilings—say $\tau_1 \leq \tau_2$ if $\zeta(\tau_1) \leq \zeta(\tau_2)$.

Theorem 4.10 *For any $4m \times 4n$ rectangle Γ, the poset P_Γ consisting of all tilings of Γ, with this order relation, is a distributive lattice.*

Proof: In order to prove that P_Γ is a lattice, we need to show that for height functions f_1 and f_2, there exists a unique greatest lower bound ("meet") α and least upper bound ("join") β. We define $\alpha(x) = \min\{f_1(x), f_2(x)\}$ and $\beta(x) = \max\{f_1(x), f_2(x)\}$, for all x. Clearly $\alpha \leq f_1$ and $\alpha \leq f_2$, and all other lower bounds are less than

α. It just remains to be shown that α is a valid height function. Clearly the values of α on the boundary will be 0, and the type-A0 points will be even and the type-A1 points will be odd, because these properties hold for f_1 and f_2. As for adjacent values differing by at most 1, suppose x and y are adjacent points, and $\alpha(x) \geq \alpha(y) + 2$. Without loss of generality, assume $\alpha(y) = f_1(y)$. Then it would follow that $f_1(x) \geq \alpha(x) \geq \alpha(y) + 2 = f_1(y) + 2$, a contradiction. Therefore, α is a valid height function. The proof for β is analogous.

To prove that P_Γ is a distributive lattice, we need to verify the distributive laws: For height functions f, g, and h,

$$(f \vee g) \wedge (f \vee h) = f \vee (g \wedge h) \qquad \text{and} \qquad (f \wedge g) \vee (f \wedge h) = f \wedge (g \vee h).$$

For any x we have:

$$((f \vee g) \wedge (f \vee h))(x) = \min(\max(f(x), g(x)), \max(f(x), h(x))).$$

The functions min and max satisfy the distributive laws, so we have

$$\min(\max(f(x), g(x)), \max(f(x), h(x))) = \max(f(x), \min(g(x), h(x))) = (f \vee (g \wedge h))(x).$$

Hence

$$(f \vee g) \wedge (f \vee h) = f \vee (g \wedge h),$$

as desired. Note that changing the sign of these functions switches the role of $\vee$ and $\wedge$, which implies the second distributive law. Therefore, P_Γ is a distributive lattice. ■

4.7 Tile invariants

In chapter 3, we computed the tile invariants for T-tetrominoes for the set of simply-connected regions $\mathcal{R}_{sc}$. We let

$$\begin{aligned} b_1(\tau) &= a_1(\tau) + a_2(\tau) + a_3(\tau) + a_4(\tau) \\ b_2(\tau) &= a_1(\tau) + a_2(\tau) \\ b_3(\tau) &= a_1(\tau) + a_4(\tau) \\ b_4(\tau) &= a_1(\tau). \end{aligned}$$

(Recall that t_1, t_2, t_3, and t_4 are the tiles which point up, right, down, and left, respectively.) Our result was that $b_1(\tau)$ is invariant mod ∞, $b_2(\tau)$ is invariant mod 4, and b_3 is invariant mod 4.

Given our local connectivity result for rectangles, we can now easily say what the tile invariants are in the case of $\mathcal{R}_{rect}$.

Theorem 4.11 *For regions in $\mathcal{R}_{rect}$, we have that $b_1(\tau)$ is invariant mod ∞, $b_2(\tau)$ is invariant mod ∞, and $b_3(\tau)$ is invariant mod ∞.*

Proof: We just need to check that b_2 and b_3 remain constant when we perform a local move. They do. ■

To prove that these invariants are the best possible, we just need to exhibit a rectangle Γ and two tilings τ_1 and τ_2 such that the number of occurrences of t_1 differs by 1. Figure 4-12 shows such an example.

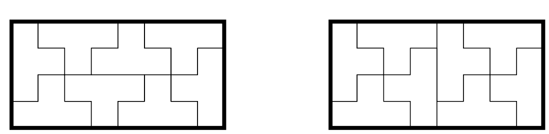


Figure 4-12: Two tilings of a rectangle which prove that b_4 may vary freely.

4.8 Non-rectangular regions

A quadruplicated simply connected region is a region which is formed by taking a simply-connected union of grid squares and dilating the figure by 4 in each direction. Let $\mathcal{Q}$ denote the set of all such regions. As we did for rectangles, we will assume that the corners of such a shape have coordinates which are congruent to (0,0) mod 4. Notice that $\mathcal{Q}$ contains all $4m \times 4n$ rectangles.

Theorem 4.12 *The second part of Theorem 4.2 holds for all regions* $\Gamma \in \mathcal{Q}$.

Proof: Suppose there exists a region $\Gamma \in \mathcal{Q}$ which can be tiled in a way which violates some of the supposed cuts and cornerless points. Let Γ' be the smallest $4m \times 4n$ rectangle which contains Γ. We can extend the tiling of Γ to a tiling of Γ' by adding tiled 4×4 squares to the part of Γ' which is not in Γ. This gives a tiling of Γ' which violates the necessary cuts and cornerless points, which contradicts Theorem 4.2. ■

As a result of this, all the above results for rectangles are also true for all $\Gamma \in \mathcal{Q}$. The proofs are the same as before.

The results do not hold if we drop the condition of being simply-connected. (Notice that the correspondence between chain graphs and height functions breaks down if the region is not simply connected, because points on the boundary of the region need not be on the unbounded face of the chain graph, so they may have nonzero height.) For example, Figure 4-13 shows a tiling of a non-simply connected region where neither the 2-move nor the 4-move can be applied.

Theorem 4.13 *The set of T-tetrominoes does not have a local-move property for* $\mathcal{R}_{sc}$.

Proof: Let Δ_1 denote the region shown in Figure 4-14. It is straightforward to see that this region can be tiled in only two ways, namely the way shown and its mirror image. Since there are no intermediate tilings, and no tile is in the same place in both tilings, the only way for local connectivity to hold for this region is if we declare this entire transformation to be one local move.

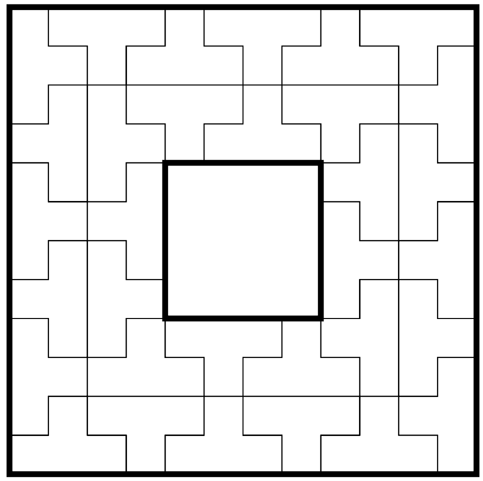


Figure 4-13: Tiling of a non-simply connected region.

In fact, we can generate infinitely many regions which admit only two tilings. Let Δ_k denote the region in Figure 4-15, where the total length of the region is $8k + 2$. As before, it can only be tiled in two ways, so in order to have local connectivity, the entire region must be considered to be a local move. No finite set of local moves can contain all of these, hence any finite set of local moves is insufficient to give local connectivity for these regions. ■

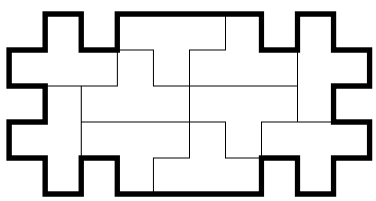


Figure 4-14: The region Δ_1.

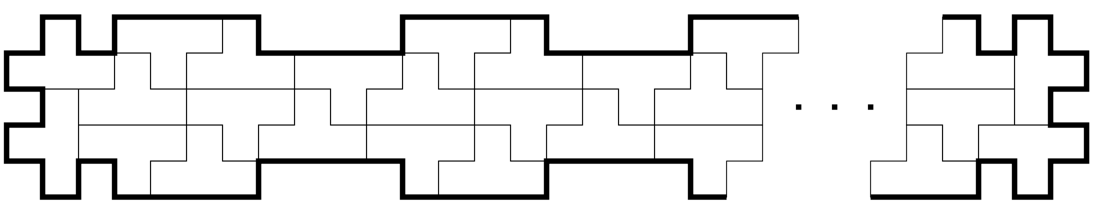


Figure 4-15: The region Δ_k.

4.9 Enumeration of tilings and the Tutte polynomial

For a region $\Gamma \in \mathcal{Q}$, define the graph G_Γ as follows. Include a vertex for each type-A1 point, and connect two vertices with an undirected edge if they are 4 units apart (vertically or horizontally). Similarly, define G^*_Γ by including a vertex for every type-A0 point, and again connecting those vertices which are 4 units apart. Note that when Γ is a $4m \times 4n$ rectangle, the graphs G_Γ and G^*_Γ are isomorphic to the $m \times n$ and $(m+1) \times (n+1)$ rectangular shape subgraphs of the square grid.

For a graph G, we let $V(G)$ and $E(G)$ denote the set of vertices and edges of G respectively. Let $c(G)$ denote the number of connected components of G. If $e \in E(G)$, let $G \backslash e$ be the graph formed by deleting e from G. Similarly, let G/e be the graph formed by contracting e in G.

The Tutte polynomial $T(G; x, y)$ is a polynomial in the variables x and y which is defined for undirected graphs G. Typically it is defined in terms of the following recursive formulas (see [29]):

- $T(G; x, y) = 1$ if G has no edges,
- $T(G; x, y) = y \cdot T(G \backslash e; x, y)$ if e is a loop,
- $T(G; x, y) = x \cdot T(G/e; x, y)$ if e is a bridge,
- $T(G; x, y) = T(G \backslash e; x, y) + T(G/e; x, y)$ if e is neither a loop nor a bridge.

Another equivalent definition of $T(G; x, y)$ is as follows. Let H be a spanning subgraph of G (that is, a subgraph of G which contains all the vertices of G). Then

$$T(G; x, y) = \sum_{H \subset G} (x-1)^{c(H)-c(G)} (y-1)^{c(H)+|E(H)|-|V(G)|}$$

where the sum is over all spanning subgraphs $H \subset G$.

Theorem 4.14 *For every $\Gamma \in \mathcal{Q}$, the number of T-tetromino tilings of Γ is equal to $2 \cdot T(G_\Gamma; 3, 3)$.*

To prove this, we introduce a few lemmas about spanning subgraphs of G_Γ and G^*_Γ.

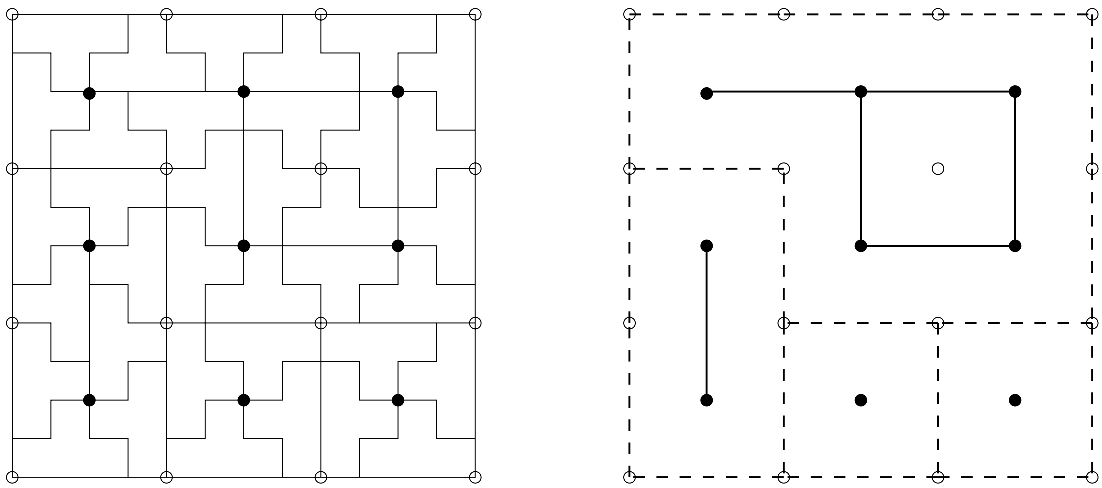


Figure 4-16: A tiling τ, and the graphs $\sigma(\tau)$ (solid lines) and $\sigma^*(\tau)$ (dotted lines).

Given a tiling τ of Γ, define $\sigma(\tau)$ to be the spanning subgraph of G_Γ which includes those edges which do not cross any tile. Similarly, define $\sigma^*(\tau)$ to be the spanning subgraph of G^*_Γ which includes those edges which do not cross any tile (see Figure 4-16).

Suppose H is a spanning subgraph of G_Γ. Define $\omega(H)$ to be the spanning subgraph of G^*_Γ consisting of those edges which do not cross any edge of H.

Lemma 4.15 *Fix $\Gamma \in \mathcal{Q}$ and a tiling $\tau \in \mathcal{Y}_\Gamma$. Then $\omega(\sigma(\tau)) = \sigma^*(\tau)$. Furthermore, no edge of the chain graph $\phi(\tau)$ crosses an edge of either $\sigma(\tau)$ or $\sigma^*(\tau)$. Conversely, any edge of G_Γ or G^*_Γ which does not cross any edge of $\phi(\tau)$ is an edge of $\sigma(\tau)$ or $\sigma^*(\tau)$.*

Proof: Notice that the points where an edge of G_Γ and an edge of G^*_Γ intersect are precisely the type-B points in the interior of Γ. Consider any such point. Recalling Figure 4-5, observe that exactly one of the two edges which meet there will avoid crossing tiles of τ. Hence each such point is on an edge of either $\sigma(\tau)$ or $\sigma^*(\tau)$, but not both. So an edge of G^*_Γ is in $\sigma^*(\tau)$ if and only if no edge of $\sigma(\tau)$ crosses it. Hence $\sigma^*(\tau) = \omega(\sigma(\tau))$.

Recall that in $\phi(\tau)$, each edge corresponds to a tile; the edge connects the two blocks in which the tile lies. Edges of $\sigma(\tau)$ and $\sigma^*(\tau)$ run along block boundaries; an edge is present in these graphs if and only if no tile crosses that boundary. If no tile crosses that boundary, then no edge of $\phi(\tau)$ will either. Conversely, if no edge of $\phi(\tau)$ crosses a block boundary, then no tile crosses that boundary, hence that boundary will be an edge of $\sigma(\tau)$ or $\sigma^*(\tau)$. (See Figure 4-17.) ■

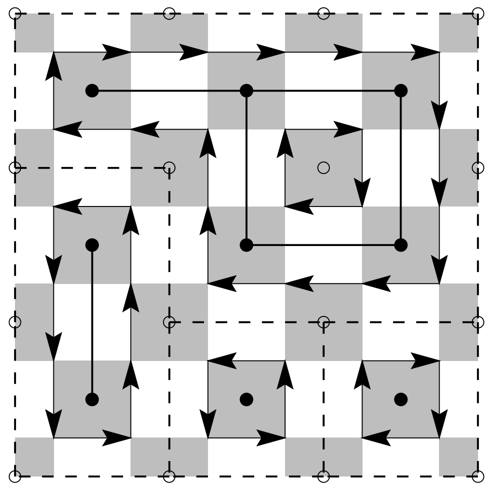


Figure 4-17: The graphs $\sigma(\tau)$ and $\sigma^*(\tau)$, and the chain graph $\phi(\tau)$.

Corollary 4.16 *Suppose a region* $\Gamma \in \mathcal{Q}$ *and tilings* $\tau_1, \tau_2 \in \mathcal{Y}_\Gamma$ *satisfy* $\sigma(\tau_1) = \sigma(\tau_2)$*. Then* $\phi(\tau_1)$ *and* $\phi(\tau_2)$ *are identical up to the orientation of the edges.*

Proof: Let $H = \sigma(\tau_1) = \sigma(\tau_2)$. For each white antiblock, there is exactly one edge of G_Γ which crosses it. The presence or absence of that edge in H determines which pair of edges along the white antiblock must be included in the corresponding chain graphs. This gives all the edges of the chain graphs, except those which do not border a complete white antiblock (ones near the boundary of the region). By inspection, one can see that all those edges must be included in order to have total degree 2 at each vertex of the chain graphs. ■

Lemma 4.17 *Let* $\Gamma \in \mathcal{Q}$*, and let* H *be a spanning subgraph of* G_Γ*. Then*

$$c(\omega(H)) = c(H) + |E(H)| - |V(G_\Gamma)| + 1.$$

Proof: We fix Γ and prove this by induction on the number of edges in H. If H has no edges, then $c(H) = |V(G_\Gamma)|$, so $c(H) + |E(H)| - |V(G_\Gamma)| + 1 = 1$, which is equal to $c(\omega(H))$, as required. Now assume that the result holds for all subgraphs $H \subset G_\Gamma$ with $|E(H)| < k$.

Consider a subgraph H with $|E(H)| = k$, and let $e \in E(H)$. First, suppose that e is a bridge of H. Then $c(H \backslash e) = c(H) + 1$, $|E(H \backslash e)| = |E(H)| - 1$, and $c(\omega(H \backslash e)) = c(\omega(H))$. We conclude:

$$\begin{aligned} c(\omega(H)) &= c(\omega(H \backslash e)) \\ &= c(H \backslash e) + |E(H \backslash e)| - |V(G_\Gamma)| + 1 \\ &= c(H) + |E(H)| - |V(G_\Gamma)| + 1. \end{aligned}$$

Now suppose that e is not a bridge of H. Then $c(H \backslash e) = c(H)$, $|E(H \backslash e)| = |E(H)| - 1$, and $c(\omega(H \backslash e)) = c(\omega(H)) - 1$. We have

$$\begin{aligned} c(\omega(H)) &= c(\omega(H \backslash e)) + 1 \\ &= c(H \backslash e) + |E(H \backslash e)| - |V(G_\Gamma)| + 2 \\ &= c(H) + |E(H)| - |V(G_\Gamma)| + 1, \end{aligned}$$

as desired. Therefore $c(\omega(H)) = c(H) + |E(H)| - |V(G_\Gamma)| + 1$ holds for all subgraphs $H \subset G_\Gamma$. ■

Suppose H is a spanning subgraph of G_Γ. Define $a(H) = 2c(H) + |E(H)| - |V(G_\Gamma)|$. Theorem 4.14 now follows from the following lemma.

Lemma 4.18 *Let Γ be a region in $\mathcal{Q}$. For every spanning subgraph $H \subset G_\Gamma$, there are exactly $2^{a(H)}$ tilings τ for which $\sigma(\tau) = H$.*

Proof: We need to show that for every spanning subgraph $H \subset G_\Gamma$, the corresponding (undirected) chain graph consists of $a(H)$ cycles. Each cycle can be oriented in two ways, hence we will get $2^{a(H)}$ valid chain graphs which correspond to H. Since chain graphs are in one-to-one correspondence with tilings, the result will follow.

Let C be a chain graph which corresponds to H. If C consists of k cycles, then it divides the plane into $k+1$ zones (possibly having holes). Each such zone is a maximal connected region on which the height function f is constant. Each zone must contain at least one type-A point, and thus must contain at least one vertex of H or $\omega(H)$. It cannot contain points from both H and $\omega(H)$, since the value of f is odd on the vertices of H and it is even on the vertices of $\omega(H)$. Observe that all vertices of H or $\omega(H)$ which live in the same zone are connected. Hence H and $\omega(H)$ have a total of $k+1$ connected components. Then $k = c(H) + c(\omega(H)) - 1 = 2c(H) + |E(H)| - |V(G_\Gamma)| = a(H)$, so the number of cycles in C is equal to $a(H)$, which proves the lemma. ■

4.10 Sampling of tilings

Let $\Gamma \in \mathcal{Q}$ be a quadruplicated simply-connected region. Define a Markov chain $\mathcal{M}$ whose states are T-tetromino tilings of Γ. Allow a transition from τ_1 to τ_2 if τ_1 and τ_2 differ by a 2-move or 4-move, with the probability of such a transition being $1/N$, where $N = |\Gamma|$ is the area of Γ. Observe that $N/2$ is larger than the maximum number of different local moves which can be applied to any one tiling. Now, let the probability of staying put in the state τ_1 be $1 - k/N \geq 1/2$, where k is the number of different local moves which can be applied to τ_1.

Observe that $\mathcal{M}$ is symmetric, and aperiodic since the probability of staying put is always $\geq 1/2$. Therefore, by Theorem 4.1, the Markov chain $\mathcal{M}$ is ergodic and converges to the uniform distribution on $\mathcal{Y}_\Gamma$. The mixing time of $\mathcal{M}$ remains open, but we would like to make the following conjecture:

Conjecture 4.19 *The mixing time of the Markov chain $\mathcal{M}$ is polynomial in the area of Γ.*

We refer the reader to [1] for the various definitions of the mixing time of Markov chains and related results. Now, if the conjecture is true, we can use the Markov chain $\mathcal{M}$ to sample tilings $\tau \in \mathcal{Y}_\Gamma$ from a nearly uniform distribution. Using the notion of self-reducibility (see introduction, [25]), we can use sampling to approximate $|\mathcal{Y}_\Gamma|$. The self-reducibility of tilings follows from the following lemma.

Lemma 4.20 *Let $\Gamma \in \mathcal{Q}$, and consider a tiling $\tau \in \mathcal{Y}_\Gamma$ chosen uniformly at random. Let S be the leftmost 4-by-4 square in the top row of Γ. Unless S is all of Γ, the probability that S is isolated in τ (covered by exactly 4 tiles) is at least 1/3 and at most 2/3.*

Proof: The 4-by-4 square S corresponds to a vertex s in G_Γ. Notice that because there is nothing to the left of S or above it, the vertex s must have degree 1 or 2 in G_Γ. The square S will be isolated if and only if no edge of $\sigma(\tau)$ is incident to s.

Case 1: Suppose s has degree 1 in G_Γ. Let e be the edge of G_Γ incident to s. Let H be a spanning subgraph of $G_\Gamma - \{s\}$. Let H_0 be the spanning subgraph of G_Γ which consists of just those edges in H, and let H_1 be the spanning subgraph of G_Γ which consists of those edges in H, plus e. Consider all tilings τ such that $\sigma(\tau)$ is either H_0 or H_1. We want to know what proportion of these tilings have $\sigma(\tau) = H_0$. Notice that $|E(H_0)| = |E(H_1)| - 1$, and $c(H_0) = c(H_1) + 1$. It follows that $a(H_0) = a(H_1) + 1$. So by Lemma 4.18, there will be twice as many tilings with $\sigma(\tau) = H_0$ as there are with $\sigma(\tau) = H_1$. This is true for any $H \subset G_\Gamma - \{s\}$. So upon picking a random tiling τ, the probability that e is present in $\sigma(\tau)$ is 1/3. So in this case, S is isolated with probability 2/3.

Case 2: Suppose s has degree 2 in G_Γ. Let e_1 and e_2 be the edges of G_Γ incident to s, and let t_1 and t_2 be the vertices adjacent to s along edges e_1 and e_2 respectively. Let H be a spanning subgraph of $G_\Gamma - \{s\}$. Let H_0 be the spanning subgraph of G_Γ which consists of just those edges in H, let H_1 be the graph which includes the edges of H plus e_1, let H_2 include the edges of H plus e_2, and let H_3 include the edges of H plus e_1 and e_2. Consider two subcases.

Subcase 2a: Suppose t_1 and t_2 are in different components of H. Notice that $|E(H_0)| = |E(H_1)| - 1 = |E(H_2)| - 1 = |E(H_3)| - 2$, and $c(H_0) = c(H_1) + 1 = c(H_2) + 1 = c(H_3) + 2$. So $a(H_0) = a(H_1) + 1 = a(H_2) + 1 = a(H_3) + 2$. So among all tilings τ which come from one of these graphs, 4/9 of them will have $\sigma(\tau) = H_0$, 2/9 of them will have $\sigma(\tau) = H_1$, 2/9 of them will have $\sigma(\tau) = H_2$, and 1/9 of them will have $\sigma(\tau) = H_3$.

Subcase 2b: Suppose t_1 and t_2 are in the same component of H. In this case, $|E(H_0)| = |E(H_1)| - 1 = |E(H_2)| - 1 = |E(H_3)| - 2$, and $c(H_0) = c(H_1) + 1 = c(H_2) + 1 = c(H_3) + 1$. So $a(H_0) = a(H_1) + 1 = a(H_2) + 1 = a(H_3)$. So among all tilings τ which come from one of these graphs, 1/3 of them will have $\sigma(\tau) = H_0$, 1/6 of them will have $\sigma(\tau) = H_1$, 1/6 of them will have $\sigma(\tau) = H_2$, and 1/3 of them will have $\sigma(\tau) = H_3$.

Combining subcases 2a and 2b, we get the following. For any H, either 1/3 or 4/9 of the tilings which correspond to H will have S isolated. Hence when we sum over all possible graphs H, we find that between 1/3 and 4/9 of all tilings of Γ have S isolated, when s has degree 2 in G_Γ.

This proves the lemma. ■

4.11 Ice graphs

Ice graphs are another type of directed graph which can be associated with a tiling. These graphs, and their associated height functions, provide another means of proving local connectivity for regions $\Gamma \in \mathcal{Q}$.

For a region $\Gamma \in \mathcal{Q}$, let B_Γ be the set of type-B points in Γ or $\partial\Gamma$. A directed graph on B_Γ is called an *ice graph* if it satisfies the following conditions:

- every two points which lie at opposite corners of the same block of Γ are connected with an edge, either one direction or the other, but not both, and
- every vertex has equal indegree and outdegree.

This notion has been explored by Eloranta [11] and others.

Let $\mathcal{I}_\Gamma$ denote the set of all ice graphs of a region Γ. Call a vertex *alternating* if it is incident to four edges which are oriented "in, out, in, out", in alternating order. Let $z(G)$ be the number of alternating vertices in an ice graph G.

In [16] the Makarychev brothers constructed a map $\mu : \mathcal{Y}_\Gamma \to \mathcal{I}_\Gamma$ as follows.

For a tiling $\tau \in \mathcal{Y}_\Gamma$, define a directed graph on B_Γ as follows. Observe that within each block, three squares belong to one T-tetromino, while one square, call it the *oddball*, belongs to a different T-tetromino. By inspection, we see that the oddball must be incident to a type-B point, rather than a type-A point. For each block, include a directed edge from the point next to the oddball square to the opposite corner of the block (see Figure 4-18). Define $\mu(\tau)$ to be the directed graph which results.

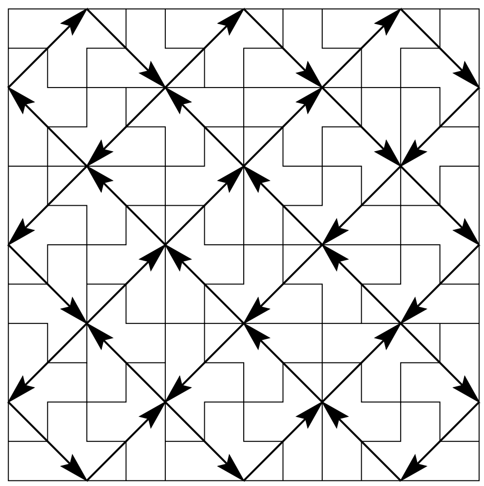


Figure 4-18: A tiling τ, and the ice graph $\mu(\tau)$.

Lemma 4.21 (K. and Y. Makarychev) *For any region $\Gamma \in \mathcal{Q}$, the map μ is a surjection from $\mathcal{Y}_\Gamma$ to $\mathcal{I}_\Gamma$, in which every ice graph G is the image of $2^{z(G)}$ tilings.*

Sketch of proof: First let us show that $\mu(\tau)$ is an ice graph. Every edge connects two opposite corners of some block, so this graph will have edges in the correct places. Notice that each type-B point is adjacent to exactly two oddballs (recall Figure 4-5), unless the point is on $\partial\Gamma$, in which case it is adjacent to only one. Therefore, every vertex has equal indegree and outdegree. So $\mu(\tau)$ is in fact an ice graph.

Now we just need to show that every ice graph G comes from exactly $2^{z(G)}$ tilings. Take a vertex of G. If the vertex is on $\partial\Gamma$, there is only one way to place the tile which touches this vertex (see Figure 4-19). Similarly, if the vertex is not on the boundary, and not alternating, there is only one way to place the two tiles which touch this

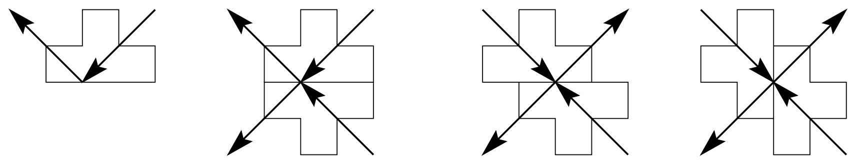


Figure 4-19: A boundary vertex, a nonalternating vertex, and the two options for an alternating vertex.

vertex. However, if the vertex is alternating, there are two ways to place the tiles around the vertex. The squares covered by the two tiles are the same in either case, so the decision of which one to use does not affect the rest of the tiling. Hence there are $2^{z(G)}$ ways to convert an ice graph G into a tiling. ■

Lemma 4.22 *If $\tau_1, \tau_2 \in \mathcal{Y}_\Gamma$ are tilings such that $\mu(\tau_1) = \mu(\tau_2)$, then τ_1 and τ_2 are local-move equivalent.*

Sketch of proof: As we just saw, the only way in which these tilings may differ is in the way the tiles next to alternating points are arranged. Converting one such configuration into the other is done by performing a 2-move. Each tile is adjacent to only one type-B point, so these moves are disjoint and can be done independently of each other. So one can convert any such tiling into any other by a sequence of 2-moves. ■

4.11.1 Height on the ice graph

For a region $\Gamma \in \mathcal{Q}$, let A_Γ be the set of type-A points in Γ or $\partial\Gamma$. Say that a function $f : A_\Gamma \to \mathbb{Z}$ is an *ice-height function* if it satisfies the following conditions:

- $f(x) = 0$ for all points $x \in \partial\Gamma$, and
- $|f(x) - f(y)| = 1$ whenever x and y are adjacent (differ by 2 in each coordinate).

Let $\mathcal{J}_\Gamma$ denote the set of all ice-height functions of a region Γ.

Theorem 4.23 *For any region $\Gamma \in \mathcal{Q}$, we have $|\mathcal{J}_\Gamma| = |\mathcal{I}_\Gamma|$.*

We define a map $\nu : \mathcal{I}_\Gamma \to \mathcal{J}_\Gamma$ as follows. Let $G \in \mathcal{I}_\Gamma$ be an ice graph. Define a function f° on the faces of G by the following rules. Let f° have the value 0 on the unbounded face of G. As we pass an edge of the graph, if the edge is oriented left-to-right as we pass it, let the value of f° increase by 1. (Similarly, if the edge is oriented right-to-left, let the value of f° decrease by 1.) Now define $f : A_\Gamma \to \mathbb{Z}$ by letting $f(x)$ equal the value of f° on the face in which x lies (see Figure 4-20). Define $\nu(G)$ to be this function f.

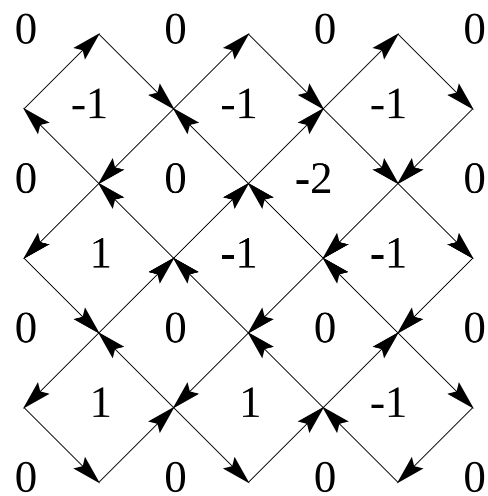


Figure 4-20: An ice graph G, and the function $f = \nu(G)$.

Theorem 4.23 will follow from the following lemma.

Lemma 4.24 *For any region $\Gamma \in \mathcal{Q}$, the map ν is a bijection between $\mathcal{I}_\Gamma$ and $\mathcal{J}_\Gamma$.*

Proof: Let G be an ice graph, and let f be $\nu(C)$. The function f is well-defined for the same reason that the height function for the chain graph is well-defined—because every vertex has equal indegree and outdegree. It is clear that such a function meets the criteria for being an ice-height function.

From an ice-height function f, one can reconstruct the ice graph $G = \nu^{-1}(f)$ by directing every edge so the face with greater height is on the left. Since the net change in height going around any vertex is 0, every vertex will have equal indegree and outdegree, thus the graph so constructed will be a valid ice graph. ■

For ease of notation, define $\xi(\tau) = \nu(\mu(\tau))$. For a region $\Gamma \in \mathcal{Q}$, the map ξ is the canonical bijection between $\mathcal{Y}_\Gamma$ and $\mathcal{J}_\Gamma$.

Lemma 4.25 *Let $\Gamma \in \mathcal{Q}$ and let $\tau_1, \tau_2 \in \mathcal{Y}_\Gamma$ be tilings of Γ. If the tilings τ_1 and τ_2 differ by a 2-move, then $\xi(\tau_1) = \xi(\tau_2)$. If the tilings τ_1 and τ_2 differ by a 4-move, then $\xi(\tau_1)$ and $\xi(\tau_2)$ differ by 2 on some point, and are the same everywhere else. If f_1 and f_2 are ice-height functions which differ by 2 on some point and are the same everywhere else, then there exist tilings τ_1 and τ_2 such that $\xi(\tau_1) = f_1, \xi(\tau_2) = f_2$, and τ_1 and τ_2 differ by a 4-move.*

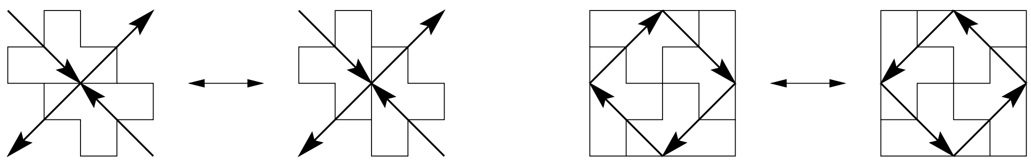


Figure 4-21: The effect of local moves on the ice graph.

Sketch of proof: A 2-move can only occur at an alternating type-B point, so if τ_1 and τ_2 differ by a 2-move, then $\mu(\tau_1) = \mu(\tau_2)$, so $\xi(\tau_1) = \xi(\tau_2)$ (see Figure 4-21).

If τ_1 and τ_2 differ by a 4-move, then $\mu(\tau_1)$ and $\mu(\tau_2)$ differ by the reversal of a directed 4-cycle, thus $\xi(\tau_1)$ and $\xi(\tau_2)$ will differ by 2 on the point inside that 4-cycle, and be the same everywhere else.

Now suppose f_1 and f_2 are ice-height functions such that $f_1(x) = h + 1$ and $f_2(x) = h - 1$, but $f_1 = f_2$ everywhere else. We must then have $f_1(y) = f_2(y) = h$ for the neighbors y of x. So x will be surrounded by a counterclockwise directed 4-cycle in the ice graph corresponding to f_1, and a clockwise directed 4-cycle in the ice graph corresponding to f_2. The problem is that a tiling which corresponds to f_1 may look like the left side of Figure 4-22. However, in such a case, there is always another tiling (which differs from the original by some 2-moves) such that a 4-move can be applied. ■

For ice-height functions $f_1, f_2 \in \mathcal{J}_\Gamma$, say that f_1 and f_2 differ by a 4-move if there exist tilings $\tau_1, \tau_2 \in \mathcal{Y}_\Gamma$ which differ by a 4-move such that $\xi(\tau_1) = f_1$ and $\xi(\tau_2) = f_2$. By the previous lemma, one can see that performing a 4-move on an ice-height function f is equivalent to increasing or decreasing its value by 2 at some point. Of course, such a move may only be applied if the function that results is a valid ice-height function.

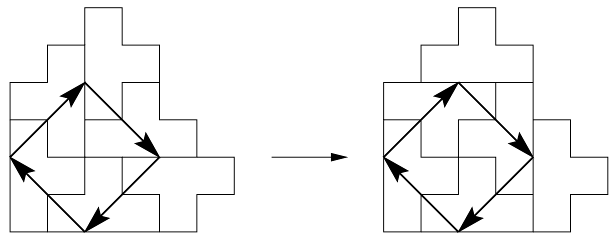


Figure 4-22: A tiling where a 4-move cannot be applied, and one where it can.

Notice that for tilings τ_1 and τ_2, having $\xi(\tau_1)$ and $\xi(\tau_2)$ differ by a 4-move does not imply that τ_1 and τ_2 differ by a 4-move. However, it does imply that there exist tilings τ_1' and τ_2' which differ by a 4-move such that $\xi(\tau_1') = \xi(\tau_1)$ and $\xi(\tau_2') = \xi(\tau_2)$. It then follows, by Lemmas 4.22 and 4.24, that τ_1 is local-move equivalent to τ_1' and τ_2 is local-move equivalent to τ_2'. Hence τ_1 and τ_2 will be local-move equivalent whenever $\xi(\tau_1)$ and $\xi(\tau_2)$ differ by a 4-move, or more generally, by a sequence of 4-moves.

Theorem 4.1 will now easily follow from the following lemma.

Lemma 4.26 *Let $\Gamma \in \mathcal{Q}$, and let $f_1, f_2 \in \mathcal{J}_\Gamma$ be ice-height functions. It is always possible to convert f_1 into f_2 by performing a sequence of 4-moves.*

Proof: For any region, there will be a unique ice-height function f_0 whose value at each point is either 0 or 1. (Each face is either "even" or "odd", depending on how many steps from the exterior it is, thus each even face will have the value 0, and each odd face will have the value 1.) It will be sufficient to show that any ice-height function f can be transformed into f_0. Suppose $f(x) > 1$ for some point x. Let x be the point where f attains its largest value, call it h (if there are several possible points, choose any one). We must then have $f(y) = h - 1$ for the neighbors y of x. Thus we can perform a 4-move, and decrease $f(x)$ to $h-2$. Repeat this process until f attains no values greater than 1. Now if there are points x where $f(x) < 0$, find the one where f attains its minimum. We can perform a 4-move to increase $f(x)$ by 2. We repeat this until f attains no values less than 0. Now $0 \leq f(x) \leq 1$ for all x, so we are done. ■

4.12 Non-quadruplicated regions

4.12.1 Cuts and cornerless points

Recall that in section 4.8, we saw an example of non-quadruplicated regions for which local connectivity does not hold. The cuts and cornerless points of Theorem 4.2 do not hold, and in general there does not seem to be any underlying structure to the set of tilings of such regions in the way that there is for quadruplicated regions.

However, for certain non-quadruplicated regions, the cuts and cornerless points will still hold. Let Γ be any simply connected tileable region. Let us say Γ is *completable* if T-tetrominoes can be added to the outside of Γ to form a rectangle. More precisely, Γ is completable if there exists a rectangle R containing Γ such that $R - \Gamma$ is tileable by T-tetrominoes.

Lemma 4.27 *If* Γ *is completable, then the appropriate cuts and cornerless points hold for tilings of* Γ*.*

Proof: The proof is the same as the proof of Theorem 4.12. Take any tiling of Γ, and extend it to a tiling of the rectangle R. This tiling must obey the cuts and cornerless points (by Theorem 4.2), hence the tiling of Γ must obey them as well. ■

There is a slight issue in determining which cuts and cornerless points are the appropriate ones. To be more precise, we should say that a region is completable only if it can be extended to a rectangle whose corners lie at points congruent to $(0, 0)$ mod 4. Of course, any region which was completable in the old sense can be translated so that it is completable in this new sense. It is the cuts and cornerless points which come from this mod-4 coordinate system which we will use.

Notice that it is possible for two distinct sets of cuts and cornerless points to hold simultaneously, as in Figure 4-23. Such cases are not particularly interesting, however, since such a region can be tiled in at most one way, as we shall see. Suppose two distinct sets of cuts and cornerless points hold. Then the arrangement must look like one of those depicted in Figure 4-24, up to rotations and reflections.

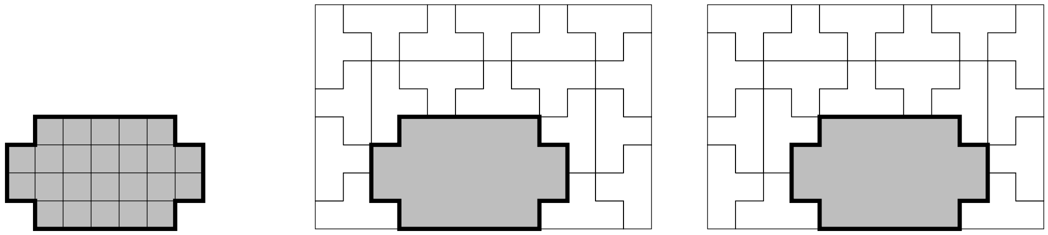


Figure 4-23: A completable region which can be extended to a rectangle in two non-equivalent ways.

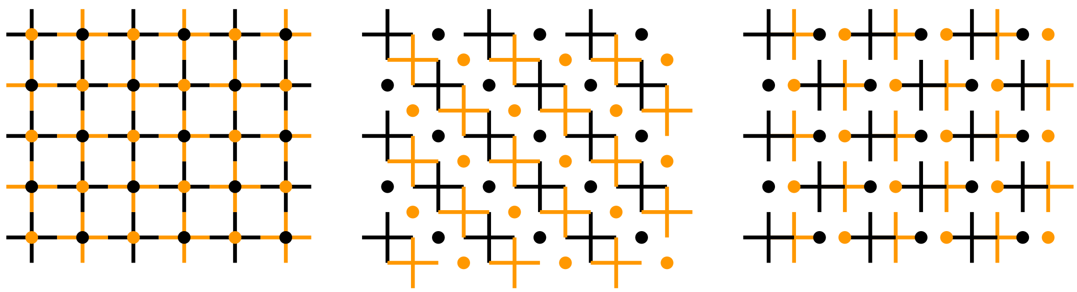


Figure 4-24: The three possibilities for overlapping sets of cuts and cornerless points. The edges are cuts; the circles are cornerless points.

In the first panel, it is clear that no tile can agree with the cuts and cornerless points, hence there can be no tilings of Γ. The same is true of the second panel; any tile which crosses no cut lines must have a corner at a cornerless point, so again there can be no tilings of Γ. With the third panel, however, it is possible to place some tiles, but there is only one way to do so. So the region Γ can be tiled in at most one way.

4.12.2 Chain graphs and height functions

Since we have cuts and cornerless points, there is some hope that we will still be able to define chain graphs and height functions. And we can, but we need to modify the definitions slightly.

Suppose Γ is a tileable completable region. Let R be a rectangle containing Γ such that $R-\Gamma$ is tileable, and let $\tau^\star$ be any tiling of $R-\Gamma$. If τ is a tiling of Γ, then the disjoint union of τ and $\tau^\star$ will be a tiling of R.

Recall that there is a one-to-one correspondence between tilings of rectangles and chain graphs, and each tile corresponds to an edge of the chain graph. Thus both $\tau^\star$ and τ will give rise to sets of directed edges, and their disjoint union will be a chain graph. Let $\phi(\tau^\star)$ be the partial chain graph which comes from the tiles in $\tau^\star$. Then it is clear that tilings of Γ will be in one-to-one correspondence with ways to complete $\phi(\tau^\star)$ to a chain graph on R.

Similarly, we can define the partial height function $f^\star = \zeta(\tau^\star)$. This assigns a height to every point of $W_{R-\Gamma}$ (recall that W_S is the set of all type-A and type-B points in S, including ∂S). Again, it is clear that tilings of Γ must be in one-to-one correspondence with ways to complete $f^\star$ to a height function of R. See Figure 4-25.

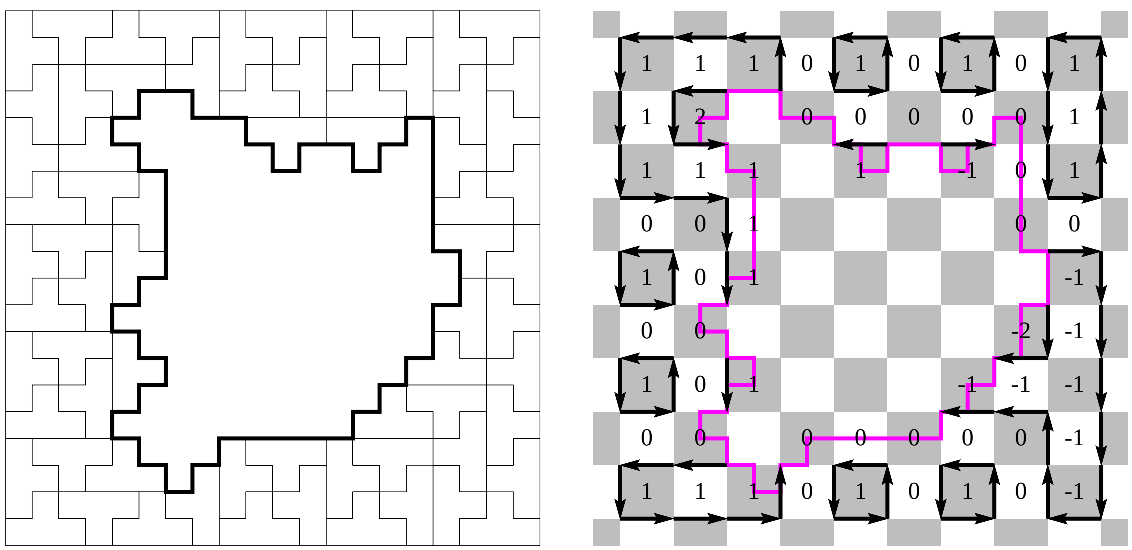


Figure 4-25: A tiling $\tau^\star$, the partial chain graph $\phi(\tau^\star)$, and the partial height function $f^\star$.

So we define a height function on Γ as follows. Say that a function $f : W_R \to \mathbb{Z}$ is a *height function on* Γ if it satisfies the following properties:

- $f(x) = f^\star(x)$ for all $x \in \partial\Gamma$ and all $x \in R - \Gamma$,
- $f(x)$ is an even integer for all type-A0 points x,
- $f(x)$ is an odd integer for all type-A1 points x, and
- $|f(x) - f(y)| \leq 1$ whenever x and y are adjacent (at a distance of two units).

As before, a local 2-move on Γ corresponds to changing the height of a type-B point by 1, and a local 4-move corresponds to changing the height of a type-A point by 2 (Lemma 4.8). Also, the set of height functions on Γ still forms a distributive lattice (Theorem 4.10). The local connectivity result (Lemma 4.9) is still true, but we must supply a different proof, since the original proof depended upon having height 0 at the boundary of Γ.

Theorem 4.28 *Let Γ be a completable region. All height functions on Γ are connected by local moves.*

Proof: The set of all height functions on Γ still forms a distributive lattice, hence there is a unique minimum height function f_0. Let f be any height function. Our goal will be to show that we can always perform a local move on f which decreases the height at a point, unless of course $f = f_0$. In this way, every height function will be local-move equivalent to f_0, proving local connectivity.

Let us write out the height function f, coloring $f(x)$ blue if $f(x) = f_0(x)$, and coloring $f(x)$ red if $f(x) > f_0(x)$ (see Figure 4-26). Observe that only points in the interior of Γ may be red. If there are no red points, then $f = f_0$, and we are done. Let k be the height of the highest red point. Assume without loss of generality that k is even.

Suppose there exists a red type-B point with height k, call it x. The type-A0 points adjacent to x must have height k as well. Let y be a type-A1 point adjacent to x. Suppose $f(y) = k + 1$. Then y cannot be red, since k is the largest red height which appears. So y must be blue, hence $f_0(y) = f(y) = k + 1$. But then $f_0(x)$ must be at least k, contradicting the fact that x is a red point. Hence we must have $f(y) = k - 1$. Thus it is possible to perform a local move to decrease the height of x.

Now we are left with the case where there is a red type-A0 point with height k, but no red type-B points of height k. Let x be such a point. If the neighbors of x all have height $k - 1$, then we can perform a local move to decrease the height of x. So suppose y is a (type-B) point adjacent to x having height at least k. It cannot be red, because all red type-B points have height at most $k - 1$. So it must be blue. Thus

$f_0(y) = f(y) \geq k$. But then $f_0(x)$ must be at least $k-1$, and since it must be even, we must have $f_0(x) \geq k$, contradicting the fact that x is red. Thus all neighbors of x have height $k-1$, so we can perform a local move to reduce the height of x.

This completes the proof. ∎

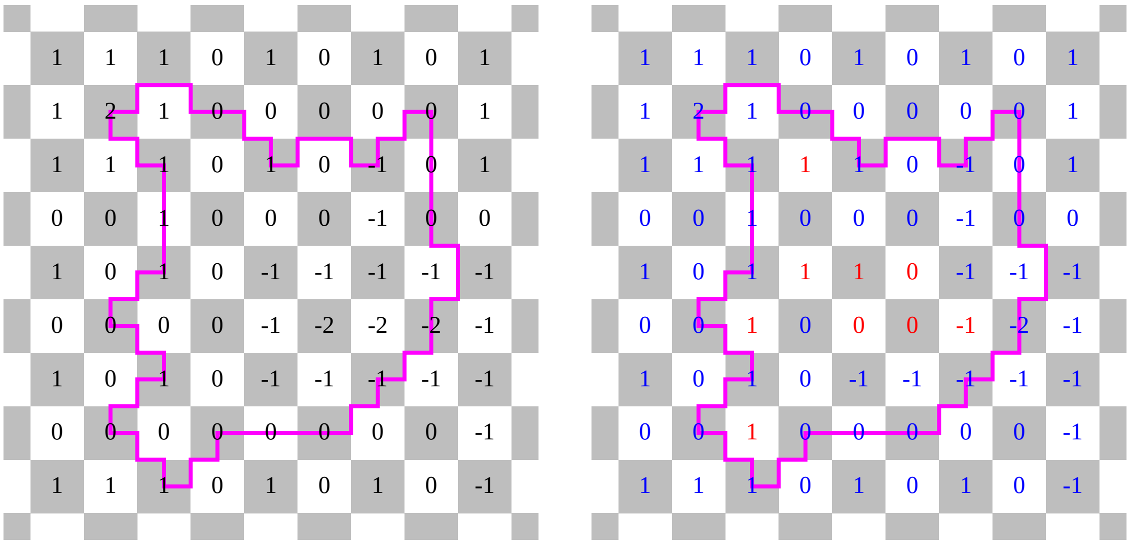


Figure 4-26: A completable region Γ. On the left is the unique minimum height function f_0. On the right is a different height function, with values colored red and blue.

4.13 The diamond-shaped completable region D_k

Let us look at a specific completable region which is particularly nice.

Consider the region shown in Figure 4-27, which we will call D_k. It is a rectangular diamond shape with $2k$ squares along one edge and $2k+1$ squares along the other edge. Notice that D_k is completable.

Now let us introduce some seemingly unrelated combinatorial objects known as *alternating-sign matrices.*

An alternating-sign matrix of order n is an $n \times n$ matrix with entries of 0, 1, and -1 such that each row and column has an odd number of non-zero entries, and these entries begin with 1 and alternate in sign. For example, the following is an alternating-sign matrix of order 7.

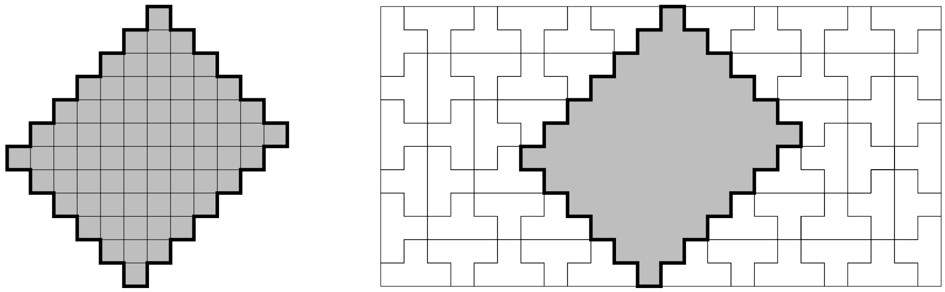


Figure 4-27: The region D_k. Here $k = 3$.

$$\begin{array}{ccccccc}
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 1 & -1 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0
\end{array}$$

Let ASM_k denote the set of all alternating-sign matrices of order k.

Alternating-sign matrices have been studied extensively in a number of previous papers, and they are related to a number of other combinatorial structures, such as the square ice model and fully-packed loop model of statistical mechanics. (See [24] for a quick overview, or [4] for an extensive history.) They also relate to T-tetromino tilings of D_k in the following way.

Theorem 4.29 *The number of T-tetromino tilings of D_k is*

$$\sum_{M \in ASM_k} 2^{\chi(M)}$$

where $\chi(M)$ denotes the number of nonzero entries in M.

Proof: Define a *saddle matrix* [1] of order k to be a $(k+1) \times (k+1)$ matrix with positive integer entries such that

[1] In [24], these are called height-function matrices. We call them saddle matrices here to avoid possible confusion with the T-tetromino height function.

- the first row and column are both $0, 1, 2, \ldots, k$,
- the last row and column are both $k, k-1, k-2, \ldots, 0$, and
- adjacent entries differ by exactly 1.

In [24], it is observed that saddle matrices are in one-to-one correspondence with ASM's. The bijection is as follows. Take a saddle matrix, and draw a filled-in circle wherever there are four neighboring entries in the form of $\begin{bmatrix} c & c+1 \\ c+1 & c \end{bmatrix}$. Likewise, draw a hollow circle wherever there are four neighboring entries in the form of $\begin{bmatrix} c+1 & c \\ c & c+1 \end{bmatrix}$. Let us call each of these arrangements a *saddle point*. Then form a $k \times k$ ASM by replacing each filled-in circle with a 1 and each hollow circle by a -1, as shown in Figure 4-28.

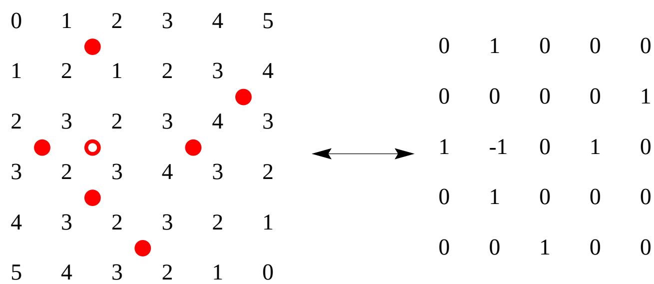


Figure 4-28: A saddle matrix, and the corresponding ASM.

Now consider tilings of D_k. They are in one-to-one correspondence with height functions on that region. Observe that taking the heights of type-A points only (and rotating the picture by 45 degrees) yields a saddle matrix (see Figure 4-29). Suppose S is a saddle matrix. Let us see what height functions correspond to it. The way to extend S to a height function is to insert the values of the type-B points between the entries of the matrix. Look at a set of four neighboring entries of S which do not form a saddle point. They must be in the form of $\begin{bmatrix} c & c+1 \\ c-1 & c \end{bmatrix}$, or some rotation of this. In this case, the type-B point which lies in the middle of these entries must have the value c. On the other hand, consider a saddle point of S. The entries must be in the form of $\begin{bmatrix} c & c+1 \\ c+1 & c \end{bmatrix}$, or a rotation of this. Now the type-B point in the middle can be either c or $c+1$. So the total number of ways to extend S to a height function is $2^{a(S)}$, where $a(S)$ is the number of saddle points in S. This is equal to $2^{\chi(M)}$, where M is the ASM corresponding to S. Thus every ASM M corresponds to $2^{\chi(M)}$ height functions (or tilings), proving the theorem. ■

Another proof of this result involves considering the ice graph corresponding to the tiling. In this case, the ice graph corresponds exactly to the "square ice" model with certain boundary conditions, which is also related to ASM's by a simple bijection.

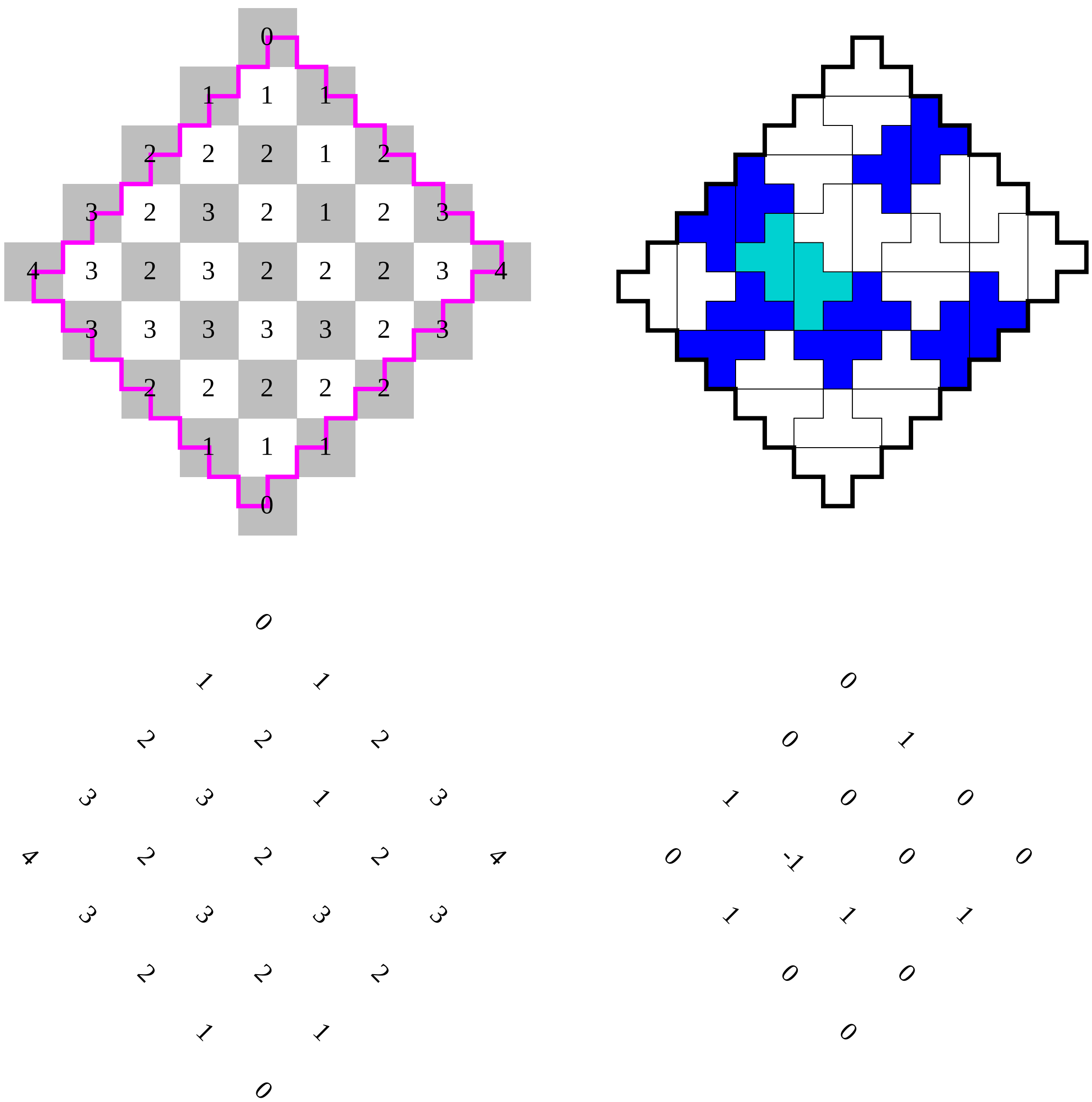


Figure 4-29: A height function on D_4, the corresponding tiling, and the associated saddle matrix and alternating-sign matrix.

For $M \in ASM_k$, let $\eta(M)$ denote the number of -1's which appear. Observe that $\chi(M) = k + 2\eta(M)$, because in each row, there is one more 1 than -1, hence the number of 1's is $\eta(M) + k$. It follows that the number of T-tetromino tilings of D_k is

$$2^k \sum_{M \in ASM_k} 4^{\eta(M)}.$$

This is interesting because there is another well-known tiling result which is very similar and has a very similar formula, namely domino tilings of Aztec diamonds (see [9] and [10]). In this case, the number of domino tilings of the order-k Aztec diamond is

$$\sum_{M \in ASM_{k+1}} 2^{\eta(M)}$$

which turns out to be equal to $2^{k(k+1)/2}$. (Unfortuantely, T-tetromino tilings of D_k do not seem to have a nice closed form like this.)

One question that has been asked about domino tilings of Aztec diamonds is the expected shape of a random tiling. Looking at a random domino tiling of a large Aztec diamond, one sees that in each corner, the tiles line up in a fixed brickwork pattern, while in some central region the tiles are essentially all mixed up. The outer regions are called the "frozen" regions, while the center is the "temperate" region. (The same property holds for ASM's; a large random ASM will generally only have non-zero entries in some temperate region in the middle.)

In this context, domino tilings are just ASM's weighted by $2^{\eta(M)}$ (and ASM's are ASM's weighted by $1^{\eta(M)}$), so it should not be too surprising that T-tetromino tilings of D_k (which are ASM's weighted by $4^{\eta(M)}$) would also exhibit this property. Figure 4-30 shows a randomly-generated T-tetromino tiling of D_{14} with tiles colored according to parity in order to highlight this phenomenon. (Tiles which point up or down are colored according to the parity of their x-coordinate, and tiles which point left or right are colored according to the parity of their y-coordinate.) Figure 4-31 shows the same tiling with local-move regions highlighted in order to show the ASM which corresponds to the tiling. (This tiling was chosen at random from an exactly uniform distribution using the technique of coupling from the past (see [21], [22], and [20]).)

It was proved in [14] that in the case of domino tilings of the Aztec diamond, the boundary of the temperate region (in the limit as $n \to \infty$) is exactly a circle. The analogous question about ASM's remains unsolved, as does the question for T-tetrominoes.

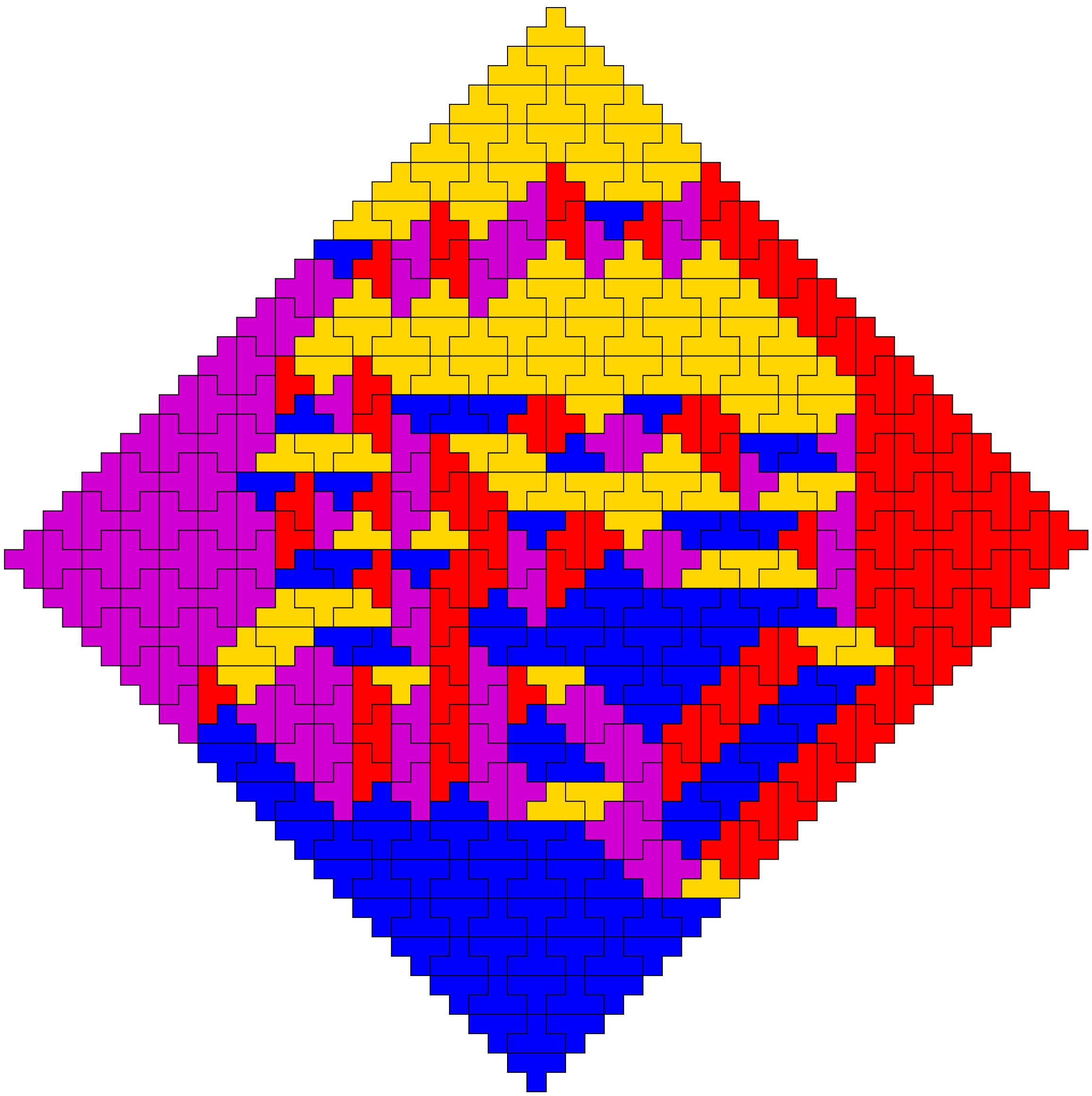


Figure 4-30: A randomly-generated tiling of D_{14}.

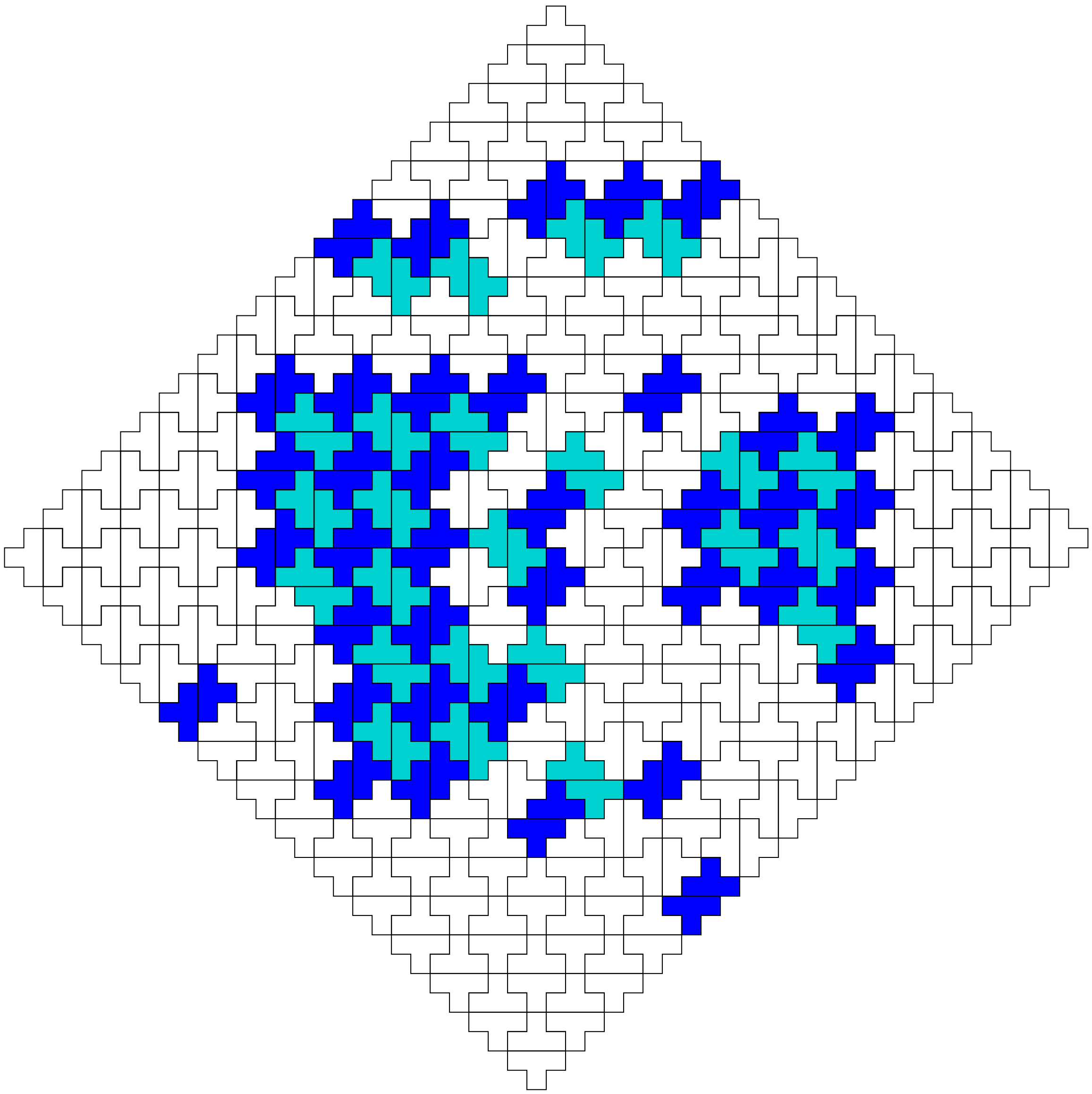


Figure 4-31: The same tiling of D_{14} with local-move regions colored.

Chapter 5

Rectangles

In this chapter we consider tiling problems where either the tiles, or the regions to be tiled, are rectangles.

5.1 Tiling rectangular regions with non-rectangular tiles

In this section, we consider the problem of tiling a rectangle with polyomino tiles. We will show that there are examples where no local move property holds.

Let $\mathcal{T}$ be the set of tiles shown in Figure 5-1. We allow rotations (but not reflections) of these tiles, so there are really 16 types of tiles. In our illustrations, each tile is colored red or gold, and is adorned with an arrow pointing in one of eight directions. We do this only to make it easier to discern the different tiles; the colors and arrows have no bearing on how the tiles can be fitted together. We will refer to tiles by their color and their arrow's direction. For instance, the first tile in Figure 5-1 would be called red-northeast.

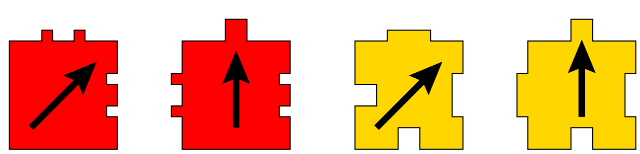


Figure 5-1: Our set of tiles $\mathcal{T}$. Rotations are allowed.

Theorem 5.1 *The set of tiles $\mathcal{T}$ has no local-move property, even if we restrict to $\mathcal{R}_{rect}$.*

Proof: Each tile is just a square with tabs or notches added to its sides. Let us scale our picture so that the underlying square has unit size. Let Γ_k be a $2k \times 2k$ square region. Figure 5-2 shows a tiling of Γ_4 using these tiles, and it is easy to see how this construction generalizes to give a tiling of Γ_k for any k. Let us call this tiling τ_0. Notice that τ_0 is not the only possible tiling of Γ_k (unless $k = 1$). For instance, Figure 5-3 shows another tiling of Γ_4, and it is easy to see how this tiling can be generalized to Γ_k.

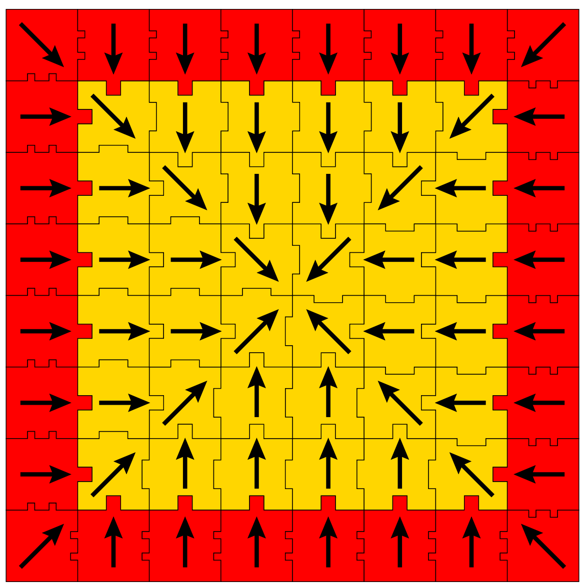


Figure 5-2: A tiling of Γ_4 with tiles from $\mathcal{T}$.

Suppose τ_1 is a tiling of Γ_k which is different from τ_0. We claim that τ_1 must differ from τ_0 in at least $2k$ places. This will imply that a local move property cannot hold. (If there were a finite set of local moves for $\mathcal{T}$, there would be some maximum size N so that every local move involved at most N tiles. Then for Γ_k with $k > N/2$, no local move could be applied to τ_0.)

Suppose the bottom b rows of τ_1 match the bottom b rows of τ_0.

First suppose $1 \leq b < k$. Then the top edge of row b consists of a double tab, followed by $b-1$ wide tabs, followed by $2k-2b$ tall tabs, followed by $b-1$ wide

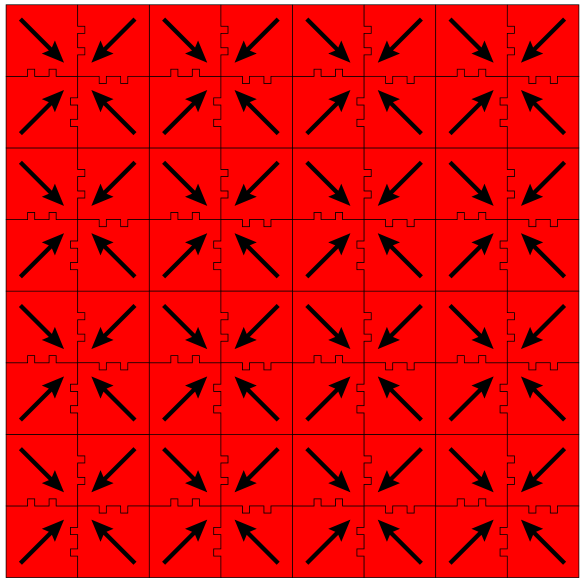


Figure 5-3: Another tiling of Γ_4 with tiles from $\mathcal{T}$.

notches, and finally a double notch. The leftmost tile of row $b+1$ must be either red-southeast or red-east. If it is red-southeast, then no tile will fit to its right. So it must be red-east. The next $b-1$ tiles sequentially can be either gold-southeast or gold-east, but gold-southeast always makes it impossible to place the next tile, so these must all be gold-east. The next tile needs a tall notch on its west and south sides, hence it must be gold-northeast. The next $2k-2b-2$ tiles sequentially can be either gold-northwest or gold-north, but gold-northwest always makes it impossible to place the next tile, so these must all be gold-north. The next tile can be either gold-northwest or gold-north, but this time, gold-north makes it impossible to place the next tile, so this one must be gold-northwest. The next $b-1$ tiles sequentially must all be gold-west. Then the final tile needs to be red-west.

So in this case, whenever the bottom b rows match those of τ_0, then row $b+1$ also matches τ_0.

Now look at the case where $k \leq b \leq 2k-2$. Now the top edge of row b consists of a double tab, followed by $2k-b-1$ wide tabs, followed by $2b-2k$ tall notches, followed by $2k-b-1$ wide notches, and finally a double notch. The leftmost tile of row $b+1$ must be either red-southeast or red-east. If it is red-southeast, then no tile will fit to its right. So it must be red-east. The next $2k-b-2$ tiles sequentially can be either gold-southeast or gold-east, but gold-southeast always makes it impossible to place the next tile, so these must all be gold-east. The next tile can be either gold-southeast or gold-east, but in this case gold-east makes it impossible to place the next tile, so this one must be gold-southeast. The next $2b - 2k$ tiles must be gold-south. The next tile must be gold-southwest. The next $2k - b - 2$ tiles must be gold-west, then the final tile must be red-west.

Lastly, if $b = 2k - 1$, then the top edge of row b consists of a double tab, followed by $2k - 2$ tall notches, followed by a double notch. So the leftmost tile of row $b + 1$ must be red-southeast. The next $2k - 2$ tiles must be red-south, and the last tile must be red-southwest.

Summarizing, if the bottom row of τ_1 is the same as the bottom row of τ_0, then by induction, all subsequent rows also have to be the same, so $\tau_1 = \tau_0$. Thus any local move that can be applied to τ_0 must involve the bottom row. By symmetry, it follows that any local move must also involve the top row (and the left and right sides, as well). Any such move must alter tiles which are a distance $2k$ from each other, hence it must involve at least $2k$ tiles. ∎

5.2 Tiling thin rectangular regions with polyomino tiles

In contrast to what we saw in the previous section, a local move property does hold if we restrict our attention to tilings of thin rectangles. This is stated more precisely in the following theorem.

Theorem 5.2 *Let $\mathcal{T}$ be any set of polyomino tiles. For an integer k, let $\mathcal{R}_k$ denote the set of rectangles having height k. Then $\mathcal{T}$ has a local move property for the set of regions $\mathcal{R}_k$.*

In order to prove this, we will need the following two lemmas about summing sets.

Define a *summing set* to be a set of positive integers S with the property that if $x \in S$ and $y \in S$, then $x + y \in S$.

Lemma 5.3 *Let S be a non-empty summing set. There exist integers N and G such that for all $x \geq N$, we have $x \in S$ if and only if x is divisible by G.*

Proof: For positive integers i, define the function $f(i)$ to be the GCD of the elements of S which are smaller than i. Notice that f is a decreasing integer-valued function, hence it has a limit (call it G). Let $S^\star$ be the set that results when we divide every element of S by G. Notice that $S^\star$ is a summing set, and the GCD of the elements of $S^\star$ is 1. So there is some finite subset of $S^\star$ whose GCD is 1.

It is well-known (see [12] for instance) that given a finite set A of integers whose GCD is 1, every sufficiently large integer can be written as a sum of elements of A (with repetitions allowed, of course). This is related to the Frobenius problem, which is to find the largest integer which cannot be written as a sum of elements of A.

At any rate, this implies that there is some integer $N^\star$ such that every integer greater than or equal to $N^\star$ is an element of $S^\star$. Hence any integer x which is divisible by G and which is at least $N = GN^\star$ is an element of S. Clearly no integer not divisible by G can be an element of S, so this proves the lemma. ■

Now let us define a slightly different type of set. Let S be a summing set. Say that a set U is an *S-summing set* if $x \in U$ and $y \in S$ implies $x + y \in U$. Notice that an S-summing set is not necessarily a summing set.

Lemma 5.4 *Let S be a non-empty summing set, and let U be a non-empty S-summing set. Let G be the GCD of the elements of S. Then there exists an integer M such that for all $x \geq M$, we have $x \in U$ if and only if $x - G \in U$.*

Proof: Let U^* denote the set of all $x \in U$ such that there does not exist $y < x$ with $y \in U$ and $y \equiv x \pmod G$. Clearly $|U^*| \leq G$. Let T be the largest member of U^*. Let N be the integer from Lemma 5.3. Let $M = T + N + G$.

Suppose $x \in U$ and $x \geq M$. There must exist some $y \in U^*$ such that $x \equiv y \pmod G$. Notice that $x - y - G \geq N$, since $x \geq M$ and $y \leq T$. Also, $x - y - G$ is divisible by G, so by Lemma 5.3, $x - y - G \in S$. Then since U is an S-summing set, $y + (x - y - G) \in U$. Thus $x \in U$ and $x \geq M$ together imply that $x - G \in U$.

The other direction is nearly identical. Suppose $x \geq M$ and $x - G \in U$. Then there exists $y \in U^*$ such that $x - G \equiv y \pmod{G}$. Like before, $x - y \geq N + G$, and $x - y$ is divisible by G, so $x - y \in S$. Then it follows that $x \in U$. ■

We are now ready to prove Theorem 5.2.

Proof: Let k be a fixed positive integer. Let A denote the set of all positive integers x such that a $k \times x$ rectangle can be tiled with tiles from $\mathcal{T}$. It is clear to see that A is a summing set. (Indeed, if a $k \times x$ rectangle and a $k \times y$ rectangle can be tiled, then these can be put next to each other to yield a tiling of a $k \times (x + y)$ rectangle.) If A is empty, then no $k \times n$ rectangle can be tiled, and Theorem 5.2 is trivially true.

Define a (k, L)-path to be a non-intersecting path along grid lines from $(0, 0)$ to $(0, k)$ which stays within the box with corners at $(0, 0)$, $(0, k)$, (L, k) and $(L, 0)$. Let P be a (k, L)-path, and let t be an integer. Define the region $\Gamma_{t,P}$ to be the shape bounded by the path P and the lines $y = 0$, $y = k$, and $x = -t$ (see Figure 5-4). For a fixed path P, let B_P be the set of all integers t such that $\Gamma_{t,P}$ is tileable by tiles from $\mathcal{T}$. Observe that B_P is an A-summing set. (Indeed, if $\Gamma_{t,P}$ is tileable, and a $k \times u$ rectangle is tileable, then the latter may be placed to the left of the former, creating a tiling of $\Gamma_{t+u,P}$.) Also observe that B may be empty. (The empty set is trivially an A-summing set.)

Figure 5-4: The region $\Gamma_{t,P}$.

Set L to be the largest horizontal length of any tile in $\mathcal{T}$. Notice that there are finitely many (k, L)-paths. For each such path P, the set B_P may or may not be empty. If B_P is non-empty, then by Lemma 5.4, there exists some integer M_P such that for all $t \geq M_P$, we have $t \in B_P$ if and only if $t - G \in B_P$ (where G is the GCD of the elements of A). Let $\widetilde{M}$ be the maximum value of M_P (where the maximum is taken over all P for which B_P is non-empty).

Let a be the smallest element of A. We are finally ready to state what our local moves are! Let our set of allowable local moves consist of all moves where the region involved fits within a $k \times (a + \widetilde{M} + L)$ rectangle. This is potentially a large number of moves, but it is certainly finite, since there are only finitely many ways that polyomino tiles can be placed within a $k \times (a + \widetilde{M} + L)$ space.

Consider a tiling of a $k \times n$ rectangle by tiles in $\mathcal{T}$. (Say that the lower left corner of this rectangle is at the origin.) If $n < a + \widetilde{M} + L$, then we are done, since any rearrangement of this rectangle is a single local move. Otherwise, draw a vertical line at $x = a + \widetilde{M}$, and color red every tile which contains some point to the left of this line (see Figure 5-5). Notice that the boundary between red and white tiles is a (k, L)-path P (translated by $a + \widetilde{M}$, of course).

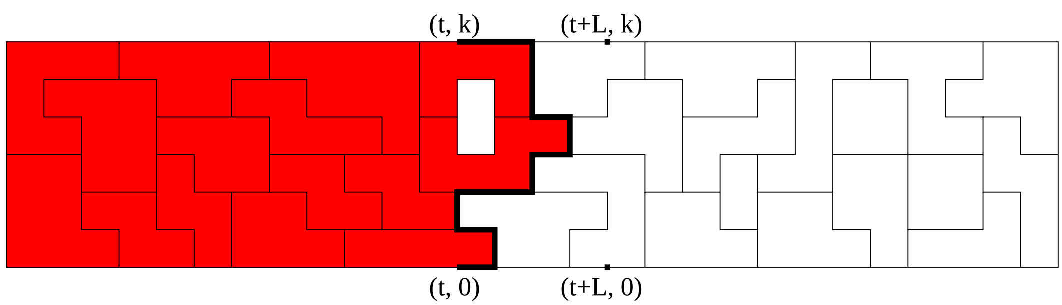


Figure 5-5: A (k, L)-path located at $t = a + \widetilde{M}$. Here $L = 4$.

Consider the set B_P. We know $a + \widetilde{M} \in B_P$, because our original tiling contains a tiling of $\Gamma_{a+\widetilde{M},P}$. Now by Lemma 5.4, we have that $a + \widetilde{M} - G \in B_P$. Applying the lemma a/G times, we get that $\widetilde{M} \in B_P$. (Recall that a is divisible by G since $a \in A$, and G is the GCD of the elements of A.) It follows that there is a tiling of $\Gamma_{a+\widetilde{M},P}$ which consists of a tiling of a $k \times a$ rectangle next to a a tiling of $\Gamma_{\widetilde{M},P}$. Since $\Gamma_{a+\widetilde{M},P}$ fits within a $k \times (a + \widetilde{M} + L)$ rectangle, we can make a local move so that our tiling now has a $k \times a$ rectangle at its left edge.

We can repeat this procedure on the $k \times (n - a)$ rectangle that remains. After several iterations, our tiling will consist of a bunch of $k \times a$ rectangles, followed by a $k \times q$ rectangle, for some $q < a + \widetilde{M} + L$. Any tiling can be converted to a tiling of this form. And all tilings of this form are local-move equivalent, since each of the subrectangles are small enough to be rearranged by a single local move. ■

5.3 Tiling with rectangular tiles

In this section we investigate the problem of tiling non-rectangular regions with rectangular tiles. If the set of allowable tiles consists of just two types of rectangles, then a local move property always holds [2]. So let us assume $\mathcal{T}$ consists of three (or more) types of rectangular tiles. The simplest non-trivial set of tiles would seem to be the following.

Consider the set of tiles $\mathcal{T}$ consisting of a 2×1 domino, a 1×2 domino, and a 1×3 tromino.

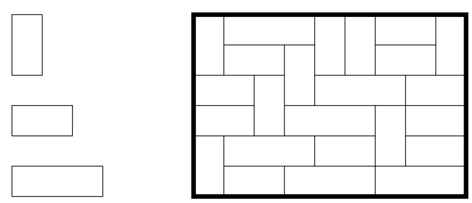


Figure 5-6: The set of tiles, and a tiling of a rectangle with these tiles.

Let $\mathcal{L}$ be the set of local moves depicted in Figure 5-7, with all orientations allowed.

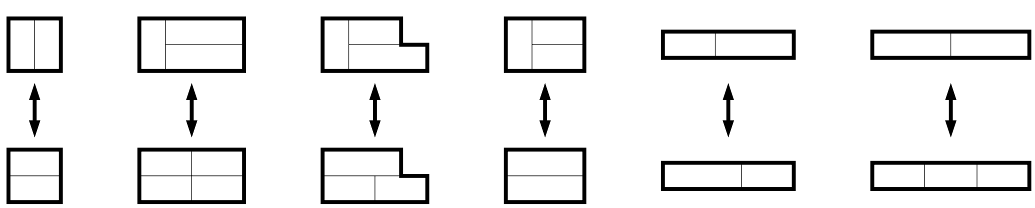


Figure 5-7: The set of local moves.

Theorem 5.5 *Every rectangular region has local connectivity with respect to $\mathcal{L}$.*

The proof of this theorem will depend on the following lemma:

Lemma 5.6 *Let Γ be an $a \times b$ rectangle, with $b \geq 2$. Let τ be a tiling of Γ which uses $k > 0$ copies of the vertical domino. Then a local move can be applied to τ to yield a tiling which uses less than k copies of the vertical domino.*

Proof: Suppose there is no local move which can be applied to τ which will reduce the number of vertical dominoes.

Let t_0 be the southernmost vertical domino in the tiling. (If there are several, any one will suffice.) Assume t_0 does not lie along the east edge of the region. (If it does, reflect the picture so that it lies along the west edge of the region.) Define the *base* of a vertical domino to be the lower of its two cells; define the *head* to be the other cell. Consider the cell c_1 directly east of the base of t_0; notice that this cell must be inside the region Γ (see Figure 5-8). The cell c_1 cannot be the head of a vertical domino, since t_0 is assumed to be the southernmost. If c_1 is the base of a vertical domino, then a local move can be applied to replace these two vertical dominoes with two horizontal dominoes.

So we may assume c_1 belongs to a horizontal tile (either a domino or a tromino). Let d_1 be the cell immediately north of c_1; notice that this cell must be in the region Γ. If d_1 belongs to a horizontal tile, then some local move can be applied which will eliminate t_0. So we may assume d_1 is the base of another vertical domino, call it t_1. Let c_2 be the cell directly east of d_1; again it must be inside the region Γ. If c_2 is the base of a vertical domino, there is a local move which can be applied. So we may assume c_2 belongs to a horizontal tile.

We may continue in this manner to create an infinite sequence of tiles moving to the northeast. Clearly this is a contradiction, proving the lemma. ∎

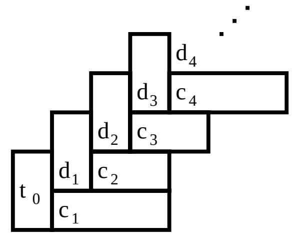


Figure 5-8: An infinite sequence of tiles in τ.

We can now prove Theorem 5.5.

Proof: If Γ has width 1, it can only be tiled in one way, and the theorem is trivial.

Assume Γ has width at least 2. By repeated use of Lemma 5.6, we see that every tiling of Γ is local-move equivalent to some tiling which uses no vertical dominoes. From any such tiling, we can use our local moves to shift all the trominoes to the right, and then convert them all into dominoes (except possibly one, if the width of Γ is odd).

Thus every tiling can be converted into this one reduced tiling, so all tilings are local-move equivalent. ∎

Notice that this local connectivity result does not generalize to all simply connected regions. For instance, the region shown in Figure 5-9 has only two tilings, so the entire switch must be considered to be a single local move.

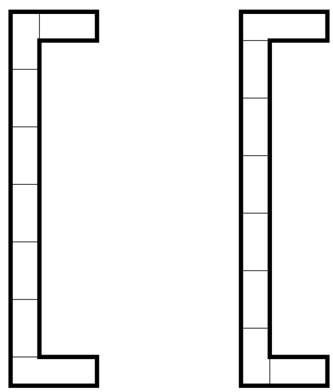


Figure 5-9: A region which is tileable in only two ways.

5.4 Open questions

There is still much that we do not know about the question of rectangle tilings. We have seen an example of three rectangular tiles which have a local-move property for $\mathcal{R}_{rect}$ but not for $\mathcal{R}_{sc}$. It seems to be pretty common for sets of (at least three) rectangular tiles not to have a local-move property for $\mathcal{R}_{sc}$. For example, let $\mathcal{T}$ be the set consisting of the 4×2, 5×3, and 3×7 rectangles, a fairly unremarkable set. Figure 5-10 shows an example of a simply connected region which admits precisely two tilings by $\mathcal{T}$. It seems that for many sets of three rectangles, a construction of this sort is possible. However, in chapter 7, we will see sets of rectangles for which local connectivity does hold for $\mathcal{R}_{sc}$.

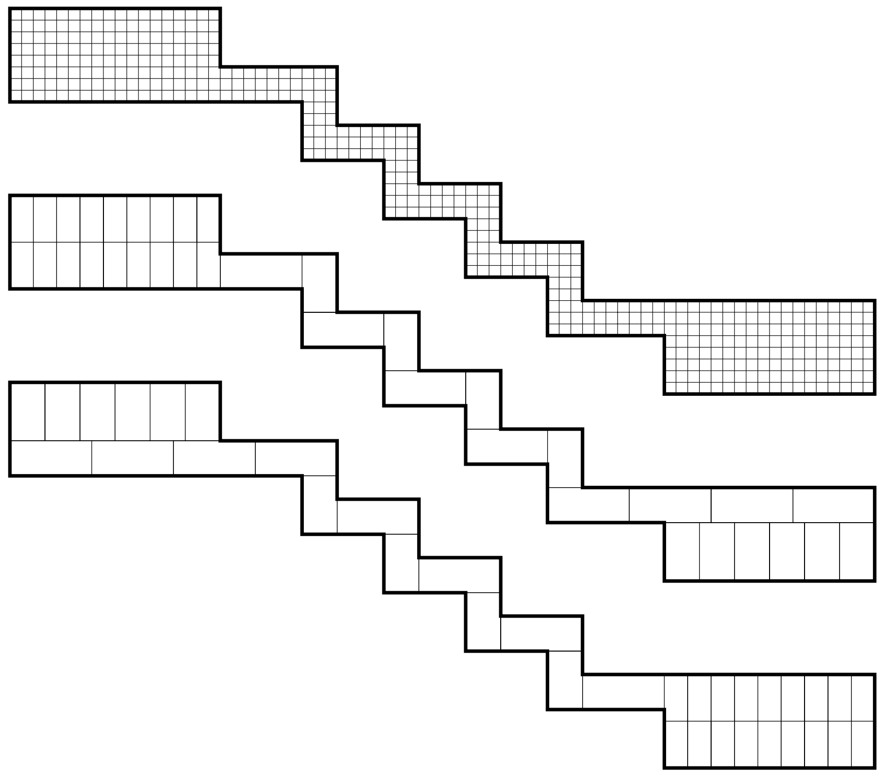


Figure 5-10: A region which admits two tilings by 4×2, 5×3, and 3×7 rectangles.

As for $\mathcal{R}_{rect}$, the problem is still open. Does every set of rectangular tiles have a local-move property for tilings of $\mathcal{R}_{rect}$?

Chapter 6

Tiling rectangles with irrational rectangles

In this chapter we consider a particular set of three rectangular tiles whose side lengths are irrational. Because these tiles are not polyominoes, some of the results we know about polyominoes will not hold for this set of tiles. In particular, we will see in section 6.7 that there is no local move property for this set of tiles, even when the region to be tiled is a rectangle of bounded height. In addition, we observe some interesting connections between tilings with these tiles and Baxter permutations.

6.1 A set of irrational tiles

Let μ and ν be positive irrational numbers.

Consider a set of tiles consisting of an $\mu \times \nu$ rectangle, a $(\mu + 1) \times 1$ rectangle, and a $1 \times (\nu + 1)$ rectangle. For convenience, we will color these rectangles purple, gold, and white respectively.

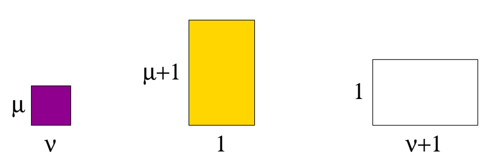


Figure 6-1: Our set of tiles. Rotations are not allowed.

Suppose Γ is a rectangle which can be tiled with these tiles. It is clear that Γ must be of the form $(p\mu + q) \times (r\nu + s)$, where p, q, r, and s are non-negative integers. In the next section we establish some necessary conditions for the coefficients p, q, r, and s.

6.2 Necessary conditions for tileability

As before, assume Γ is a $(p\mu + q) \times (r\nu + s)$ rectangle, and Γ is tileable by our set of tiles.

Suppose that one of these coefficients, say p, is 0. Then Γ will be of the form $q \times (r\nu + s)$. Since the rectangle has height q, only white tiles may be used, since a purple or gold tile would add an unwanted value of μ to the height. Such a tiling is in some sense trivial. (A similar thing happens if q, r, or s is 0.) From now on, we will only look at rectangles which have non-trivial tilings. So we must have p, q, r, and s strictly positive.

Theorem 6.1 *Let Γ be a $(p\mu + q) \times (r\nu + s)$ rectangle. Then Γ is tileable if and only if $qs = ps + qr$.*

We will prove the "only if" direction here, and prove the "if" direction in section 6.6.

Observe that Γ has area $pr \cdot \mu\nu + ps \cdot \mu + qr \cdot \nu + qs$. The areas of the tiles are $\mu\nu$, $\mu + 1$, and $\nu + 1$ respectively. If Γ admits a tiling by these tiles, it must use pr copies of the purple tile. Furthermore, it must use ps copies of the gold tile, and qr copies of the white tile. Also, qs must be the total number of gold and white tiles used. Hence we must have $qs = ps + qr$.

6.3 Tilings and permutations

For this section, we will take $p = 1$, $q = n$, $r = n - 1$, and $s = n$. Notice that this satisfies the condition in Theorem 6.1. So Γ is a $(\mu + n) \times ((n - 1)\nu + n)$ rectangle. Figure 6-2 shows a tiling of Γ.

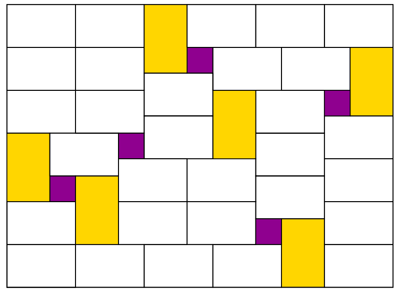
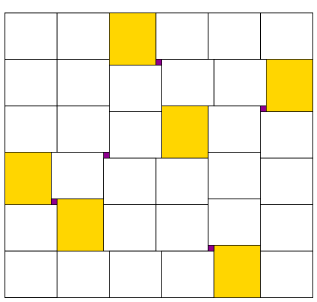


Figure 6-2: A tiling τ of Γ. Here $n = 6$.

Let $\mathcal{Y}_n$ denote the set of tilings of Γ. Let S_n denote the set of permutations of n elements (or equivalently, the set of $n \times n$ permutation matrices). We define a map $\phi : \mathcal{Y}_n \to S_n$ as follows. Begin by drawing the tiling, taking $\mu = \nu = \epsilon$, where ϵ is small, as in Figure 6-3. If we take the limit of this tiling as $\epsilon \to 0$, the purple tiles become nonexistent, and the tiling becomes an $n \times n$ grid, with some of the grid cells gold and others white. Replace each gold tile with a 1 and each white tile with a 0. Any vertical line across Γ must have length $\mu + n$, so it must cross $n - 1$ white tiles and 1 gold tile. Hence the resulting matrix must have one 1 in each column. By a similar argument, there must also be one 1 in each row. Hence this gives us a permutation matrix.



$$\begin{matrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{matrix}$$

Figure 6-3: The same tiling τ redrawn, and the matrix $\phi(\tau)$.

At first glance, one might think that every permutation matrix corresponds to a tiling. But this is not the case.

6.4 Baxter permutations

Baxter permutations are classical objects which have been studied in a number of earlier papers, such as [3], [6], and [8]. They were first introduced in connection with a problem about compositions of continuous functions, but since then have received a fair amount of attention as a purely combinatorial object.

Definition 6.2 *A permutation σ on n elements is called a Baxter permutation if there do not exist integers $i < j < k$ such that*

$$\sigma(j) < \sigma(k) < \sigma(i) < \sigma(j+1) \qquad \textit{or} \qquad \sigma(j+1) < \sigma(i) < \sigma(k) < \sigma(j).$$

For instance, the permutation 5147623 is a Baxter permutation, but 5147236 is not (take $i = 3$, $j = 4$, and $k = 7$, for instance). Let $\mathcal{B}_n$ denote the set of Baxter permutations of size n.

Now let us consider another definition of Baxter permutations. Let $\sigma \in S_n$ be a permutation. For $1 \leq x \leq n-1$ and $1 \leq y \leq n-1$, call an ordered pair (x, y) a *vortex* of σ if either

- $\sigma(x) < y < y+1 < \sigma(x+1)$ and $\sigma^{-1}(y+1) < x < x+1 < \sigma^{-1}(y)$, or
- $\sigma(x+1) < y < y+1 < \sigma(x)$ and $\sigma^{-1}(y) < x < x+1 < \sigma^{-1}(y+1)$.

Lemma 6.3 *A permutation σ is a Baxter permutation if and only if it has no vortex.*

This result was independently proved in [5] [1], but we will provide our own proof here for completeness.

Proof: Suppose (x, y) is a vortex of σ, and assume without loss of generality that

$$\sigma(x) < y < y+1 < \sigma(x+1) \quad \text{and} \quad \sigma^{-1}(y+1) < x < x+1 < \sigma^{-1}(y).$$

Now setting $i = \sigma^{-1}(y+1)$, $j = x$, and $k = \sigma^{-1}(y)$ in Definition 6.2 yields that σ is not a Baxter permutation. This show the "only if" direction.

[1] In [5], a vortex was called a *windmilled configuration.*

Now suppose σ is not a Baxter permutation, and assume without loss of generality that there exist integers $i < j < k$ such that

$$\sigma(j) < \sigma(k) < \sigma(i) < \sigma(j+1).$$

Let c be in the range $\sigma(k) \leq c \leq \sigma(i)$. Color c red if $\sigma^{-1}(c) < j$ and color c blue if $\sigma^{-1}(c) > j+1$ (observe that $\sigma^{-1}(c)$ cannot equal j or $j+1$). Notice that $\sigma(k)$ is blue and $\sigma(i)$ is red. Thus there must be some c_0 such that c_0 is blue and $c_0 + 1$ is red. Then (j, c_0) is a vortex. ∎

Returning to tilings, we have the following theorem.

Theorem 6.4 *The map ϕ is a bijection from $\mathcal{Y}_n$ to $\mathcal{B}_n$.*

Proof: Let us consider the process of converting a permutation matrix into a tiling, i.e., computing ϕ^{-1}.

Begin with an $n \times n$ permutation matrix M. As usual, we will label the rows $1, 2, \ldots, n$ from top to bottom, and label the columns $1, 2, \ldots, n$ from left to right.

From this matrix, we will attempt to create a potential tiling of Γ. Let us declare the upper-left corner of Γ to be the origin. (This may seem a little bit awkward, but unfortunately, the way the entries of a matrix are typically indexed does not correspond nicely to the way points in Cartesian coordinates are indexed.) As before, it will be helpful to think of μ and ν being small.

Each entry of M corresponds to either a white or gold tile. For $1 \leq i \leq n$ and $1 \leq j \leq n$, let $A_{i,j} = 1$ if the 1 in row i occurs to the left of column j, and let $A_{i,j} = 0$ otherwise. Similarly, let $B_{i,j} = 1$ if the 1 in column j occurs above row i, and let $B_{i,j} = 0$ otherwise. If $M = \phi(\tau)$ for some tiling τ, it is clear that the tile in row i and column j must have $A_{i,j}$ gold tiles and $j - 1 - A_{i,j}$ white tiles to its left. Similarly, it must have $B_{i,j}$ gold tiles and $i - 1 - B_{i,j}$ white tiles above it. Thus the upper-left hand corner of this tile must occur at the point with coordinates $((j - 1 - A_{i,j})\nu + (j - 1), -B_{i,j}\mu - (i - 1))$, and the tile must be gold if the entry is a 1, and white if the entry is a 0.

So for any matrix M, the exact location of all the white and gold tiles in $\phi^{-1}(M)$ is determined. This proves that ϕ is one-to-one. What remains to be shown is that $\phi^{-1}(M)$ is a valid tiling if and only if M is a Baxter permutation matrix.

It is clear from the construction that two tiles which are horizontally or vertically adjacent will not overlap each other. It is also clear that the tiles extend just to the edge of Γ and leave no gaps on the boundary. The potential problem occurs with tiles that are diagonally adjacent. Consider the four tiles which lie in rows i or $i+1$, and columns j or $j+1$. Up to rotations and reflections, the place where these tiles meet will look like one of the pictures in Figure 6-4.

Figure 6-4: The four types of corners which can appear.

In the first two instances, the tiles meet exactly and leave no gaps. In the third instance, the tiles leave an $\mu \times \nu$ gap, requiring us to place a purple tile there. In the fourth instance, however, two tiles overlap, and thus we do not get a valid tiling. It follows that the permutation matrices which generate a tiling are precisely those for which this bad case never occurs. Upon inspection, we see that the bad case corresponds exactly to the presence of a vortex. Hence by Lemma 6.3, the permutation matrices which avoid the bad case are those which are Baxter permutations. ■

6.5 Generalizing to other rectangles Γ

In the previous sections, we took Γ to be a $(p\mu + q) \times (r\nu + s)$ rectangle, with $p = 1$, $q = n$, $r = n - 1$, and $s = n$. Suppose we consider other values for p, q, r, and s (while still obeying the relation $qs = ps + qr$ from section 6.2). See Figure 6-5.

Let Γ be a $(p\mu + q) \times (r\nu + s)$ rectangle. Let $\mathcal{Y}_{p,q,r,s}$ denote the set of tilings of Γ. As before, let us take a tiling $\tau \in \mathcal{Y}_{p,q,r,s}$, and draw the tiling with $\mu = \nu = \epsilon$, and take the limit as $\epsilon \to 0$. Let us replace each gold tile with a 1, and each white tile

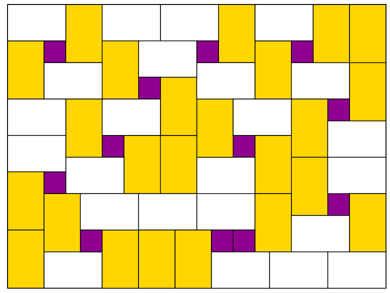
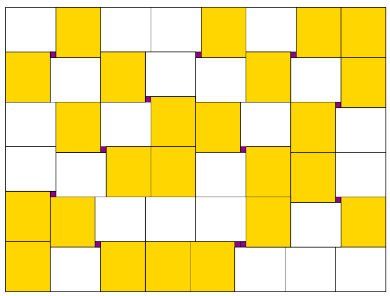


Figure 6-5: A tiling τ of a $(3\mu + 6) \times (4\nu + 8)$ rectangle.



$$\begin{array}{cccccccc}
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 0 & 0
\end{array}$$

Figure 6-6: The same tiling τ, and the matrix $\phi(\tau)$.

with a 0, and call the resulting matrix $\phi(\tau)$. The result is a $q \times s$ matrix of 0's and 1's which has r 0's and $s - r$ 1's in each row, and has $q - p$ 0's and p 1's in each column.

Let us define a *doubly-balanced matrix* to be a 0/1-matrix such that there exist integers i and j such that every row contains i 1's and every column contains j 1's. Let $\mathcal{M}_{p,q,r,s}$ denote the set of all $q \times s$ doubly-balanced matrices which have r 0's and $s - r$ 1's in each row, and have $q - p$ 0's and p 1's in each column. Notice that if we count by rows, such a matrix contains $q(s - r)$ 1's, and if we count by columns, the matrix contains ps 1's. Hence $\mathcal{M}_{p,q,r,s}$ is empty unless $qs = ps + qr$. Also notice that $\mathcal{M}_{1,n,n-1,n}$ is just the set of permutation matrices.

It is clear that ϕ is a map from $\mathcal{Y}_{p,q,r,s}$ to $\mathcal{M}_{p,q,r,s}$.

Now let us generalize the notion of a vortex (from section 6.4) so that it applies to any doubly-balanced matrix.

Let M be a $q \times s$ doubly-balanced matrix, and let $m_{i,j}$ denote its i, j-entry. Let $A_{a,b} = \sum_{j=1}^{b-1} m_{a,j}$, and let $B_{a,b} = \sum_{i=1}^{a-1} m_{i,b}$. In words, $A_{a,b}$ is the number of 1's among the entries directly to the left of the a, b-entry, and $B_{a,b}$ is the number of 1's among the entries directly above the a, b-entry.

For $1 \leq a \leq q-1$ and $1 \leq b \leq s-1$, call an ordered pair (a, b) a *vortex* of M if either

- $A_{a,b+1} < A_{a+1,b+1}$ and $B_{a+1,b} > B_{a+1,b+1}$, or
- $A_{a,b+1} > A_{a+1,b+1}$ and $B_{a+1,b} < B_{a+1,b+1}$.

Notice that when M is a permutation matrix, this notion is equivalent to the earlier definition of a vortex. Let $\mathcal{B}_{p,q,r,s}$ denote the subset of matrices in $\mathcal{M}_{p,q,r,s}$ which contain no vortex. Notice that $\mathcal{B}_{1,n,n-1,n}$ is the set of Baxter permutation matrices.

This brings us to the following generalization of Theorem 6.4.

Theorem 6.5 *The map ϕ is a bijection from $\mathcal{Y}_{p,q,r,s}$ to $\mathcal{B}_{p,q,r,s}$.*

Proof: The proof runs along the same lines as the proof of Theorem 6.4. We will begin with a matrix $M \in \mathcal{M}_{p,q,r,s}$ and try to construct the tiling $\phi^{-1}(M)$.

Each entry of M corresponds to either a white or gold tile. If $M = \phi(\tau)$ for some tiling τ, it is clear that the tile in row i and column j must have $A_{i,j}$ gold tiles and $j-1-A_{i,j}$ white tiles to its left. Similarly, it must have $B_{i,j}$ gold tiles and $i-1-B_{i,j}$ white tiles above it. Thus the upper-left hand corner of this tile must occur at the point with coordinates $((j-1-A_{i,j})\nu + (j-1), -B_{i,j}\mu - (i-1))$, and the tile must be gold if the entry is a 1, and white if the entry is a 0.

So for any matrix M, the exact location of all the white and gold tiles in $\phi^{-1}(M)$ is determined. This proves that ϕ is one-to-one. What remains to be shown is that $\phi^{-1}(M)$ is a valid tiling if and only if $M \in \mathcal{B}_{p,q,r,s}$.

It is clear from the construction that two tiles which are horizontally or vertically adjacent will not overlap each other. It is also clear that the tiles extend just to the edge of Γ and leave no gaps on the boundary. The potential problem occurs with tiles that are diagonally adjacent. Consider the four tiles which lie in rows i or $i+1$, and columns j or $j+1$. Up to rotations and reflections, the place where these tiles meet will look like one of the pictures in Figure 6-4.

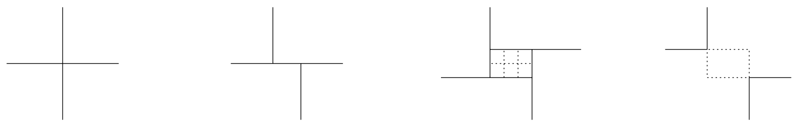


Figure 6-7: The four types of corners which can appear.

In the first two instances, the tiles meet exactly and leave no gaps. In the third instance, the tiles leave a gap of size $(B_{i+1,j}-B_{i+1,j+1})\mu\times(A_{i,j+1}-A_{i+1,j+1})\nu$, requiring us to place some purple tiles there. In the fourth instance, however, two tiles overlap, and thus we do not get a valid tiling. It follows that the permutation matrices which generate a tiling are precisely those for which this bad case never occurs. Upon inspection, we see that the bad case occurs when $A_{i,j+1} > A_{i+1,j+1}$ and $B_{i+1,j} < B_{i+1,j+1}$, or vice versa. This is exactly the definition of a vortex. Hence those matrices which generate tilings are those which have no vortex. ■

6.6 Conditions for tileability revisited

In section 6.2 we observed that in order for a $(p\mu + q) \times (r\nu + s)$ rectangle to be tileable, we must have $qs = ps + qr$. Let us now show that this condition is also sufficient.

Let Γ be a $(p\mu + q) \times (r\nu + s)$ rectangle with $qs = ps + qr$. Let $\omega = \frac{p}{q} = \frac{s-r}{s}$. Write $\omega = \frac{x}{y}$, where x and y are relatively prime. It is clear that both q and s must be divisible by y. It follows that Γ can be covered by $(x\mu + y) \times ((y - x)\nu + y)$ rectangles. So it will suffice to exhibit a tiling of a $(x\mu + y) \times ((y - x)\nu + y)$ rectangle. Equivalently, it will suffice to exhibit a $y \times y$ doubly-balanced matrix H with x 1's in each row and column such that H has no vortex.

For integers $0 < x < y$, define H as follows. Let

$$h_{i,j} = \begin{cases} 1 & \text{if} \quad i + j \leq x \\ 0 & \text{if} \quad x < i + j \leq y \\ 1 & \text{if} \quad y < i + j \leq y + x \\ 0 & \text{if} \quad y + x < i + j \end{cases}$$

Figure 6-8 shows the matrix H, and the tiling $\phi^{-1}(H)$. It is straightforward to check that H does not contain a vortex, hence $\phi^{-1}(H)$ is well-defined. Thus a $(x\mu + y) \times ((y - x)\nu + y)$ rectangle can be tiled, so therefore Γ can be tiled. This completes the proof of Theorem 6.1.

$$\begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{matrix}$$

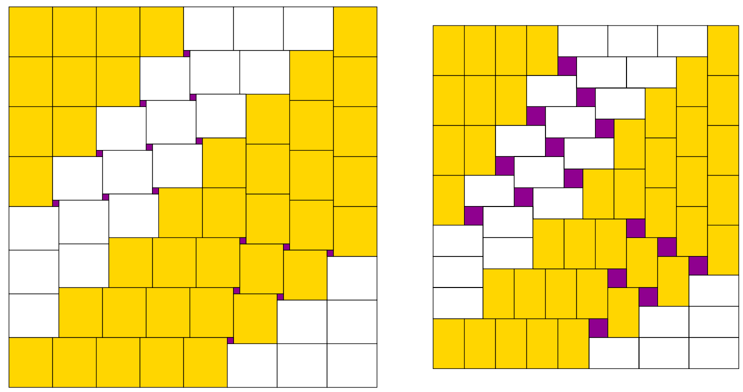


Figure 6-8: The matrix H for $x = 5$ and $y = 8$.

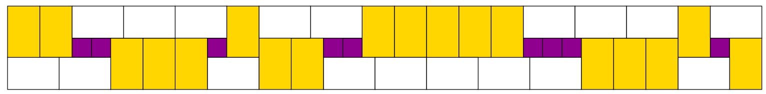
6.7 Local moves

Let us consider local moves for this set of tiles. It turns out that local connectivity fails badly in this case. In Theorem 5.2, we proved that if $\mathcal{T}$ is any set of polyomino tiles and k is any finite number, then there is a local-move property for tilings of rectangles of height k. The tiles we consider now are not polyominoes, and in fact, Theorem 5.2 does not hold.

Theorem 6.6 *This set of tiles does not have a local-move property for rectangles of height* $\mu + 2$.

Consider rectangles of the shape $(\mu + 2) \times (k\nu + 2k)$, for some integer k. Figure 6-9 shows an example of such a tiling τ. Using the ideas from section 6.5, each such tiling τ corresponds to a matrix $\phi(\tau) \in \mathcal{B}_{1,2,k,2k}$. This is a vortex-free $2 \times 2k$ matrix having one 1 in each column, and k 1's in each row.

In this case, the set of matrices $\mathcal{B}_{1,2,k,2k}$ has a particularly nice description. Let $M \in \mathcal{M}_{1,2,k,2k}$, and let $A_{i,j}$ be defined as in section 6.5. For $0 \leq c \leq 2k$, define



$$\begin{array}{cccccccccccccccccc} 1&1&0&0&0&1&0&0&1&1&1&1&1&0&0&0&1&0\\ 0&0&1&1&1&0&1&1&0&0&0&0&0&1&1&1&0&1 \end{array}$$

Figure 6-9: A tiling τ of a $(\mu+2)\times(9\nu+18)$ rectangle, and the associated matrix $\phi(\tau)$.

$\kappa(c) = A_{1,c+1} - A_{2,c+1}$. (So $\kappa(c)$ is the number of 1's in the first c columns of the top row, minus the number of 1's in the first c columns of the bottom row.) It is clear that $\kappa(c) = \kappa(c-1) \pm 1$ for $c > 0$, and that $\kappa(0) = \kappa(2k) = 0$.

Lemma 6.7 *The matrix M has a vortex if and only if the function κ has a positive local minimum or a negative local maximum.*

Proof: Suppose κ has a positive local minimum. Then for some c, $\kappa(c) > 0$ and $\kappa(c-1) = \kappa(c+1) = \kappa(c)+1$. Then the entries of the matrix around this point must be $m_{1,c} = 0$, $m_{2,c} = 1$, $m_{1,c+1} = 1$, and $m_{2,c+1} = 0$. Thus $B_{2,c} = 0$ and $B_{2,c+1} = 1$. And $\kappa(c) > 0$ implies that $A_{1,c+1} > A_{2,c+1}$. Thus $(1,c)$ is a vortex of M. A similar thing happens if κ has a negative local maximum.

Now suppose M has a vortex. Certainly it must occur at a point with coordinates $(1,c)$, with c in the range $1 \le c \le 2k-1$. Suppose it is the case that $A_{1,c+1} > A_{2,c+1}$ and $B_{2,c} < B_{2,c+1}$. The first statement implies that $\kappa(c) > 0$. Then $B_{2,c} < B_{2,c+1}$ implies that $B_{2,c} = 0$ and $B_{2,c+1} = 1$, hence $m_{1,c} = 0$ and $m_{1,c+1} = 1$. It then follows that $m_{2,c} = 1$ and $m_{2,c+1} = 0$, and thus $\kappa(c+1) = \kappa(c)+1$ and $\kappa(c-1) = \kappa(c)+1$. Thus κ has a positive local minimum. Similarly, if $A_{1,c+1} < A_{2,c+1}$ and $B_{2,c} > B_{2,c+1}$, then κ must have a negative local maximum. ■

Let $\mathcal{H}_k$ denote the set of all functions $\kappa : \{0,1,\ldots,2k\} \mapsto \mathbb{Z}$ such that $\kappa(0) = \kappa(2k) = 0$, $\kappa(c) = \kappa(c-1) \pm 1$, and such that κ has no positive local minimum or negative local maximum.

Let κ_0 be defined by $\kappa_0(c) = c$ for $0 \le c \le k$, and $\kappa_0(c) = 2k - c$ for $k \le c \le 2k$.

Lemma 6.8 *If $\kappa \in \mathcal{H}_k$ and $\kappa \ne \kappa_0$, then κ differs from κ_0 in at least $k-1$ values.*

Proof: Suppose we had $\kappa(c) > 0$ for all c in $0 < c < 2k$. Since κ is not allowed to have positive local minima, the only minima of κ must occur at the endpoints. Thus κ must consist of a monotonic increase from 0 to ts maximum, followed by a monotonic decrease to 0. The only allowable such function is κ_0.

So it must be the case that $\kappa(c) = 0$ for some c in $0 < c < 2k$. This c must be even, because consecutive values of κ always change by 1, hence they alternate parity. Consider a value d in the range $\frac{c}{2} < d < \frac{c+2k}{2}$. Observe that $\kappa(d) \leq |d - c|$, and that $\kappa_0(d)$ is either d or $2k - d$. For $\frac{c}{2} < d \leq c$, we have that $\kappa(d) \leq c - d < 2k - d$, since $c < 2k$. Also $\kappa(d) \leq c - d < d$, since $d > \frac{c}{2}$. So in either case, $\kappa(d) < \kappa_0(d)$. For $c \leq d < \frac{2k-d}{2}$, we have that $\kappa(d) \leq d - c < d$, since $c > 0$. Also $\kappa(d) \leq d - c < 2k - d$, since $d < \frac{2k-d}{2}$. Again we have $\kappa(d) < \kappa_0(d)$. Thus for all d in this range, $\kappa(d) \neq \kappa_0(d)$. There are $k - 1$ values in this range, thus proving the lemma. ∎

We are now on the brink of proving Theorem 6.6. Consider the tiling τ_0 shown in Figure 6-10. This is the tiling which corresponds to the function κ_0. Suppose τ is any other tiling of this region. The function κ which coresponds to it must differ from κ_0 in at least $k - 1$ places, by Lemma 6.8. Hence the parts of τ which differ from τ_0 must include tiles which are at a distance of at least $k - 1$ from each other. Thus for k large enough, there is no local move that can be applied to τ_0.

This proves Theorem 6.6.

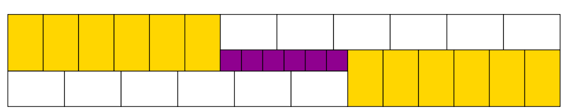


Figure 6-10: The tiling τ_0, shown here for $k = 6$.

Chapter 7

Tiling with generalized dominoes

7.1 Introduction

Define a *generalized domino of order k* to be a rectangle with integer-length sides whose area is 2^k. Notice that there are $k+1$ generalized dominoes of order k. Generalized dominoes of order 1 are the usual dominoes. Let $\mathcal{T}_k$ denote the set of generalized dominoes of order k.

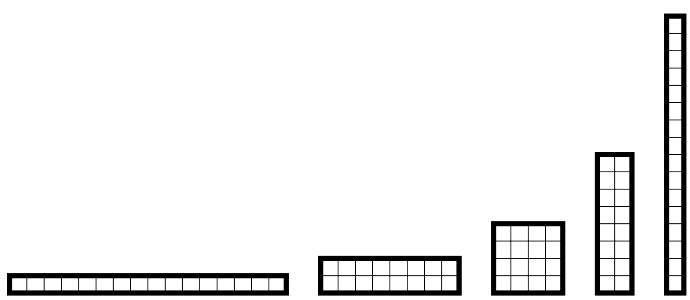


Figure 7-1: The generalized dominoes of order 4.

Define a set of local moves $\mathcal{L}_k$ as follows. For $1 \leq c \leq k$, observe that a $2^c \times 2^{k+1-c}$ rectangle can be tiled in two ways. Let it be a local move to convert one such tiling into the other. Figure 7-2 shows the four local moves for $\mathcal{T}_4$.

Our goal is to prove the following theorem.

Theorem 7.1 *Fix an integer k. Tilings of simply connected regions with the tile set $\mathcal{T}_k$ have local connectivity with respect to $\mathcal{L}_k$.*

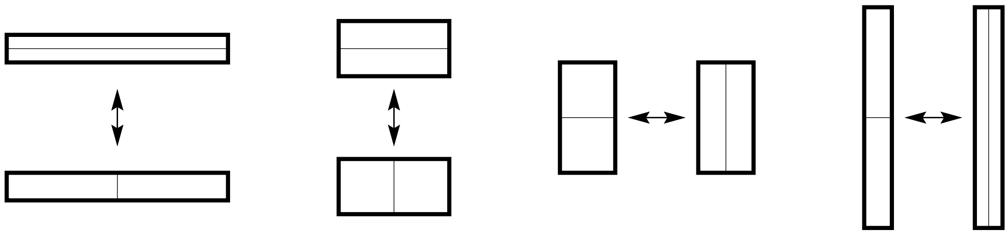


Figure 7-2: The local moves for $\mathcal{T}_4$.

7.2 The height function

For a non-zero integer x, define $\theta(x)$ to be the greatest integer c such that 2^c divides x. For instance, $\theta(16) = 4$, $\theta(12) = 2$, and $\theta(19) = 0$. (Define $\theta(0) = \infty$.) Observe the following facts about θ:

- If $\theta(x) > \theta(y)$, then $\theta(x + y) = \theta(y)$.
- If $\theta(x) = \theta(y)$, then $\theta(x + y) > \theta(y)$.
- If $x < y$ and $c \geq 0$, and $\theta(x) > c$ and $\theta(y) > c$, then there exists z in $x < z < y$ such that $\theta(z) = c$.

Fix an integer k, and let $\mathcal{W}$ denote the set of words of the form

$$a^{e_1} b^{e_2} a^{e_3} \cdots a^{e_{2m-1}}$$

or

$$a^{e_1} b^{e_2} a^{e_3} \cdots b^{e_{2m}}$$

(for any m) such that $\theta(e_i) + \theta(e_{i+1}) < k$ for all i in the proper range. We will call $\mathcal{W}$ the set of *perfect* words. For a perfect word $w \in \mathcal{W}$, define its length $|w|$ to be the number of multiplicative terms (either $2m - 1$ or $2m$, in this case).

Define an equivalence relation on $\mathcal{W}$ as follows. Let w be a perfect word. Say

$$w = \cdots b^{e_{i-1}} a^{e_i} b^{e_{i+1}} \cdots .$$

Let Q be any integer such that $\theta(Q) + \theta(e_i) \geq k$. Then we will let w be equivalent to w', where

$$w' = \cdots b^{e_{i-1}+Q} a^{e_i} b^{e_{i+1}-Q} \cdots .$$

Notice that w' is a perfect word, for the following reason. We have $\theta(e_{i-1}) + \theta(e_i) < k$, since w is perfect, thus $\theta(e_{i-1}) < \theta(Q)$, so $\theta(e_{i-1} + Q) = \theta(e_{i-1})$, by our facts about θ. Hence the θ values do not change, so the resulting word will still be perfect.

Now define the equivalence relation by the transitive closure of this operation (and the analogous operation with the role of a and b switched). Essentially this says we are allowed to push b^Q through a^{e_i}, provided $\theta(Q) + \theta(e_i) \geq k$. (And the same thing with a and b switched.)

For a perfect word $w \in \mathcal{W}$, define its *θ-profile* to be the sequence

$$\Theta(w) = (\theta(e_1), \theta(e_2), \ldots, \theta(e_{|w|})).$$

Any two consecutive terms of $\Theta(w)$ must always sum to less than k.

Lemma 7.2 *If w_1 and w_2 are equivalent perfect words, then $\Theta(w_1) = \Theta(w_2)$.*

Proof: It suffices to show this result for the case when w_1 and w_2 differ by a single operation. And we have already shown this to be true. ∎

The values that our height function will take will be equivalence classes of $\mathcal{W}$. (For convenience, we will just write each height as a single perfect word, and remember that it represents the entire equivalence class.) Also, we will call them *labels* rather than *heights*, to avoid any possible confusion. The way we will choose these words is as follows. Begin with some arbitrary point on the boundary of Γ, and assign it the label $(ab)^N$, for some sufficiently large value of N. (Taking N to be larger than the area of Γ should be sufficient. We do this because we don't want to deal with what happens when our perfect words reach length 0. A similar approach was used by Kenyon and Kenyon in [15].) Then to define labels for other points of the tiling, we walk along tile boundaries. When we move east, we right-multiply by a; when we move west, we right-multiply by a^{-1}; when we move north, we right-multiply by

b; and when we move south, we right-multiply by b^{-1}. Unfortunately, we have to be careful here, because we have not yet defined multiplication for perfect words (or equivalence classes of perfect words). So let us pause to do that.

Let $w \in \mathcal{W}$ be a perfect word, and let c be an integer. Define the product $w \cdot a^c$ as follows. If the last term of w is a power of b, append an a^c at the end. If the last term of w is a power of a, add c to the exponent. Now, if the result is a perfect word, we are done. If not, say we have $va^x b^y a^z$, where v is some expression involving a and b. Since w was perfect, and this word is not, we have $\theta(x) + \theta(y) < k$, and $\theta(y) + \theta(z) \geq k$. In this case, we will let the product $w \cdot a^c$ equal $va^{x+z}b^y$. Notice that $\theta(z) > \theta(x)$, hence $\theta(x+z) = \theta(x)$, so this is perfect. We define $w \cdot b^c$ similarly.

First, we must show this is well-defined on equivalence classes. In other words, we have to show that if w_1 is equivalent to w_2, then $w_1 \cdot a^c$ is equivalent to $w_2 \cdot a^c$. It suffices to show this for the case where w_1 and w_2 differ by a single operation. The only thing that affects if an operation can be done is the θ-value of the middle exponent. Upon multiplying by a^c, the only θ that can change is the last one (it can also disappear). The only case that needs to be checked is pushing a power of a past the last power of b. It is straightforward to verify that things work in this case.

Also, it is necessary to verify that $(w \cdot a^{c_1}) \cdot a^{c_2}$ is equivalent to $w \cdot a^{c_1+c_2}$. We leave this as an exercise for the reader.

All that remains to show is that upon completing a loop around a tile, the resulting word is equivalent to the one we began with. This is the key step in defining any height function of this sort, and is the whole reason why we have to define our words in precisely the way we do.

Lemma 7.3 *Let w be a perfect word. Fix p in the range $0 \leq p \leq k$, and let $U = 2^p$ and let $V = 2^{k-p}$. Then w is equivalent to $wa^U b^V a^{-U} b^{-V}$ (where multiplication takes place from left-to-right).*

Proof: We consider six cases.

Case 1: $w = va^x b^y a^z$, and $\theta(z) < p$.

Then $wa^U = va^x b^y a^{z+U}$, which is perfect since $\theta(z+U) = \theta(z)$.

Then $wa^U b^V = va^x b^y a^{z+U} b^V$, which is perfect since $\theta(z+U) < p$.

Then $wa^U b^V a^{-U} = va^x b^y a^{z+U} b^V a^{-U}$, which becomes $va^x b^y a^z b^V$.

Then $wa^U b^V a^{-U} b^{-V} = va^x b^y a^z$, as desired.

Case 2: $w = va^x b^y a^z$, $\theta(z) \geq p$, and $\theta(z+U) + \theta(y) < k$.

Then $wa^U = va^x b^y a^{z+U}$, which is perfect since $\theta(z+U) + \theta(y) < k$.

Then $wa^U b^V = va^x b^y a^{z+U} b^V$, which becomes $va^x b^{y+V} a^{z+U}$.

Then $wa^U b^V a^{-U} = va^x b^{y+V} a^z$, which is perfect, since $\theta(y) < k - p$, so $\theta(y+V) = \theta(y)$.

Then $wa^U b^V a^{-U} b^{-V} = va^x b^{y+V} a^z b^{-V}$, which becomes $va^x b^y a^z$, as desired.

Case 3: $w = va^x b^y a^z$, $\theta(z) = p$, and $\theta(z+U) + \theta(y) \geq k$.

Then $wa^U = va^x b^y a^{z+U}$, which becomes $va^{x+z+U} b^y$.

Then $wa^U b^V = va^{x+z+U} b^{y+V}$, which is perfect, since $\theta(y) < k - p$, so $\theta(y+V) = \theta(y)$.

Then $wa^U b^V a^{-U} = va^{x+z+U} b^{y+V} a^{-U}$, which is perfect since $\theta(y+V) < k - p$.

Then $wa^U b^V a^{-U} b^{-V} = va^{x+z+U} b^{y+V} a^{-U} b^{-V}$, which becomes $va^{x+z+U} b^y a^{-U}$. This is equivalent to w since $\theta(z+U) + \theta(y) \geq k$.

Case 4: $w = vb^x a^y b^z$, and $\theta(z) < k - p$.

Then $wa^U = vb^x a^y b^z a^U$, which is perfect since $\theta(z) < k - p$.

Then $wa^U b^V = vb^x a^y b^z a^U b^V$, which becomes $vb^x a^y b^{z+V} a^U$.

Then $wa^U b^V a^{-U} = vb^x a^y b^{z+V}$.

Then $wa^U b^V a^{-U} b^{-V} = vb^x a^y b^z$, as desired.

Case 5: $w = vb^x a^y b^z$, $\theta(z) \geq k - p$, and $\theta(z+V) + \theta(y) < k$.

Then $wa^U = vb^x a^y b^z a^U$, which becomes $vb^x a^{y+U} b^z$.

Then $wa^U b^V = vb^x a^{y+U} b^{z+V}$, which is perfect since $\theta(y+U) = \theta(y) < k - \theta(z+V)$.

Then $wa^U b^V a^{-U} = vb^x a^{y+U} b^{z+V} a^{-U}$, which becomes $vb^x a^y b^{z+V}$.

Then $wa^U b^V a^{-U} b^{-V} = vb^x a^y b^z$, as desired.

Case 6: $w = vb^x a^y b^z$, $\theta(z) = k - p$, and $\theta(z+V) + \theta(y) \geq k$.

Then $wa^U = vb^x a^y b^z a^U$, which becomes $vb^x a^{y+U} b^z$.

Then $wa^Ub^V = vb^xa^{y+U}b^{z+V}$, which becomes $vb^{x+z+V}a^{y+U}$.

Then $wa^Ub^Va^{-U} = vb^{x+z+V}a^y$, which is perfect since $\theta(x+z+V) = \theta(x)$.

Then $wa^Ub^Va^{-U}b^{-V} = vb^{x+z+V}a^yb^{-V}$, which is equivalent to w, because $\theta(z+V) + \theta(y) \geq k$.

This proves the lemma. ■

We have finally shown that this height function is well-defined.

7.3 The size function

Now let us define a size function on elements of $\mathcal{W}$. The size $S(w)$ will be an ordered pair of integers. If $w \in \mathcal{W}$, let $S(w)$ be the ordered pair (m, c), where $m = |w|$ is the number of multiplicative terms, and $c = \theta(e_{|w|})$ is the θ-value of the final exponent. For example, if $k \geq 2$, $S(a^{15}b^{13}a^9b^3a^{18}b^{15}a^{14})$ would be the ordered pair $(7, 1)$. Notice that the size of w can be determined just by knowing $\Theta(w)$. Hence by Lemma 7.2, the size function is invariant on equivalence classes.

If w_1 and w_2 are elements of $\mathcal{W}$, then we will say w_1 is larger than w_2 if $S(w_1)$ is lexicographically greater than $S(w_2)$. Specifically, w_1 is larger than w_2 if w_1 has more multiplicative terms than w_2, or if they have the same number of terms, but the final exponent of w_1 is divisible by a higher power of 2 than the final exponent of w_2.

The proof of local connectivity will follow from the following lemma.

Lemma 7.4 *Fix a value k, and let τ be a tiling of a simply connected region Γ with tiles from $\mathcal{T}_k$. Define the height function as before. Let P be the point in the tiling with the largest label, and let its size be (m, c). Let t be a tile which touches P.*

If m is odd, then P must be the midpoint of the top or bottom edge of t, and the dimensions of t must be $2^{k-c-1} \times 2^{c+1}$.

If m is even, then P must be the midpoint of the left or right edge of t, and the dimensions of t must be $2^{c+1} \times 2^{k-c-1}$.

Proof: Without loss of generality, let m be odd. The case where m is even is identical, except with the x- and y-coordinates switched.

Let the label of P be $va^xb^ya^z$, where v is some expression involving a and b. Since the size of this label is (m, c), we must have $\theta(z) = c$. Let $\theta(y) = d$. Notice that $c + d < k$, or else this would not be perfect. Also we must have $\theta(x) < k - d$.

Let us consider the tile boundaries in the vicinity of P. There cannot be edges heading north or south from P, since this would make the label of a neighboring point larger than the label of P. So P must lie along a horizontal tile boundary. Without loss of generality, let us say that the tile t lies above P.

Let A and B denote the lower-left and lower-right corners of t, respectively. Let the distance from P to A be $z - r$, so the label of A becomes $va^xb^ya^r$. Since there is an edge heading north from A, $va^xb^ya^r$ cannot be perfect, or else the point north of A would have a larger label than P. So we must have $\theta(r) \geq k - d$. Thus the label of A must actually be $va^{x+r}b^y$. Similarly, let the distance from P to B be $s - z$, so the label of B becomes $va^xb^ya^s$. Then we must have $\theta(s) \geq k - d$, and the label of B would actually be $wa^{x+s}b^y$.

Since r and s are both divisible by 2^{k-d}, it follows that the width of t (which equals $s - r$) must be at least 2^{k-d}. By one of the properties of θ, we know there exists a value u satisfying $r < u < s$ such that $\theta(u) = k - d - 1$. Thus there is some point along the line from A to B whose label is $va^xb^ya^u$ (notice that this is necessarily perfect). The size of this label is $(m, k - d - 1)$. We assumed that (m, c) was the largest size achieved, hence we must have $c \geq k - d - 1$. We observed earlier that $c + d < k$, hence we must have $c = k - d - 1$.

Let h be the vertical dimension of t. We claim that $h = 2^d$. Suppose, to the contrary, that $\theta(h) < d$. Let C and D be the upper-left and upper-right corners of t, respectively. The label of C would be $va^{x+r}b^{y+h}$, and the label of D would be $va^{x+s}b^{y+h}$. (Notice that these are perfect, since $\theta(x+r) = \theta(x)$ and $\theta(y+h) = \theta(h) < d$.) The area of t is 2^k, so $h \cdot (s - r) = 2^k$, thus $\theta(s - r) = k - \theta(h) > k - d$. Consider the midpoint of the top edge of t. Its label will be $va^{x+r}b^{y+h}a^{(s-r)/2}$. Observe that $\theta((s-r)/2) + \theta(y+h) = \theta(s-r) - 1 + \theta(h) = k - 1$, hence this expression is perfect. And $\theta((s-r)/2) = \theta(s-r) - 1 > k - d - 1 = c$, so the size of this label is greater than that of P, a contradiction. So we must have $\theta(h) \geq d$. But $s - r \geq 2^{k-d}$, so we must have $h = 2^d$ and $s - r = 2^{k-d}$. So in terms of c, we have $h = 2^{k-c-1}$ and $s - r = 2^{c+1}$. So the dimensions of t must be $2^{k-c-1} \times 2^{c+1}$.

Notice that r and s differ by 2^{c+1}, and both r and s are divisible by 2^{c+1}. Meanwhile, z is divisible by 2^c. It follows that z must equal $(s + r)/2$, so P must be the midpoint of segment AB. ■

7.4 Local connectivity

We can now prove Theorem 7.1.

Proof: This proof will be very similar to Thurston's proof of local connectivity for domino tilings [27]. We will show that if the largest label does not occur on the boundary, then there is a local move that can be applied which will reduce the size of the largest label. And if the largest label does occur on the boundary, then there is a specific tile which is forced into place, and hence we may consider the smaller region Γ' which has that tile removed from it.

We will use induction on the size of Γ. Obviously if the size of Γ is 2^k, then there can be only one tile, hence only one tiling, so local connectivity holds trivially. So assume Γ is larger than this, and that local connectivity holds for all regions smaller than Γ.

First we will show that if the largest label does not occur on the boundary of Γ, then there is a local move which will reduce the size of the largest label.

Let τ be any tiling, and let P be the point of the tiling which has the largest label. Let the size of this label be (m, c), and assume without loss of generality that m is odd. Then P is the midpoint of the bottom edge of a $2^{k-c-1} \times 2^{c+1}$ tile, and is also the midpoint of the top edge of a $2^{k-c-1} \times 2^{c+1}$ tile. These two tiles put together form a $2^{k-c} \times 2^{c+1}$ rectangle, so a local move can be applied. Let us call this rectangle Ω.

We are replacing the horizontal edge which bisected Ω with a vertical one. This eliminates the point which had the largest label. Now we just need to check that the points on this new vertical edge have labels which are strictly smaller. Let E be the midpoint of the right edge of Ω, let F be the upper-right corner of Ω, and let G be the midpoint of the top edge of Ω.

Let us say, as before, that the label of P is $va^zb^ya^z$. We know that $\theta(z) = c$, and from the proof of Lemma 7.4, we have that $\theta(y) = k - c - 1$. The label of E is then $va^xb^ya^{z+2^c}$. Since $\theta(z) = \theta(2^c) = c$, we have that $\theta(z + 2^c) \geq c + 1$, so this expression should be reduced to $va^{x+z+2^c}b^y$. Then the label of F would be $va^{x+z+2^c}b^{y+2^{k-c-1}}$. This may or may not be perfect. Assume for now that it is perfect.

Continuing around Ω, we have that the label of G must be $va^{x+z+2^c}b^{y+2^{k-c-1}}a^{-2^c}$. We have that $\theta(y) = \theta(2^{k-c-1}) = k - c - 1$, hence $\theta(y + 2^{k-c-1}) \geq k - c$. And $\theta(-2^c) = c$, so the label of G is actually $va^{x+z}b^{y+2^{k-c-1}}$. Any point lying on the new vertical edge will have a label of the form $va^{x+z}b^{y+2^{k-c-1}-q}$ for some q. The number of multiplicative terms here is at most $m - 1$, hence this label is strictly smaller than the label of P was.

On the other hand, if the label of F were not perfect, then it would reduce to a word of length $m - 2$ which ends with a power of a. Then the label of G would be another word of length $m - 2$ which ends in a power of a, or possibly an even shorter word, if it reduces again. In any event, the labels of points lying on the new vertical edge will have length at most $m - 1$, so these labels will still be strictly smaller than the label of P was.

Hence this local move eliminates the largest label, and introduces only labels which are strictly smaller. Repeated application of such moves will continue to reduce the size of the largest label, until the largest label is one that is on the boundary. Hence any tiling of Γ is local-move equivalent to one which has its largest label on the boundary.

What remains to be shown is that all tilings of Γ whose largest label lies on the boundary are local-move equivalent. Let τ_1 and τ_2 be two such tilings. The labels on the boundary of Γ do not depend on the tiling, hence this largest label is the same for each tiling. Say the point is P, and its height is (m, c), where again we assume without loss of generality that m is odd. Then by Lemma 7.4, we know that in both τ_1 and τ_2, there is a $2^{k-c-1} \times 2^{c+1}$ tile with P at the midpoint of its top or bottom edge. This tile is the same in both tilings, so let Γ' be the region which remains when we remove this tile from Γ. Let τ_1' and τ_2' be the tilings obtained by deleting this tile from τ_1 and τ_2. It is possible that Γ' may be disconnected, but in any case, Γ' consists of one or more simply-connected regions. By induction, local connectivity holds for tilings of Γ'. Thus τ_1' and τ_2' are local-move equivalent, so τ_1 and τ_2 must be local-move equivalent as well.

This proves the theorem. ■

7.5 Final remarks

We have proved that for the generalized dominoes of order k, a small set of local moves is sufficient to give local connectivity for tilings of simply-connected regions.

Observe that for the case of $k = 1$, this reduces to the standard domino height function approach. Every perfect word is equivalent to one of the form $a^{m_1}b^{m_2}abababa\cdots$, and the total power of a represents the x-coordinate of the point, and the total power of b represents the y-coordinate of the point. Thus the only information that is interesting in such a word is its length, which is exactly equivalent to the standard height.

There is certainly the notion of local moves which make the tiling "higher" and "lower", but it is not immediately clear whether one can define a lattice of tilings like in the case of dominoes. What does it mean to take the "minimum" of two tilings?

Chapter 8

Tiling with polyominoes of height 2

8.1 Introduction

In [18], Pak introduces a set of tiles which consists of two horizontal T-tetrominoes and two horizontal skew-tetrominoes. He calls this the *two-row set*. Observe that this set can tile an infinite strip of height 2, and that there is a natural set of local moves which provides local connectivity.

We will consider this set of tiles, as well as a similar set of tiles where the skew-tetrominoes are replaced by a horizontal domino. This is perhaps a natural modification to make, since a skew-tetromino can be tiled by two horizontal dominoes. However, as we shall see, these tile sets are quite different in terms of local connectivity and tile invariants.

8.2 Horizontal T-tetrominoes and horizontal skew-tetrominoes

Let $\mathcal{T}$ be the set of tiles consisting of the horizontal T-tetrominoes and the horizontal skew-tetrominoes, as shown in Figure 8-1. Figure 8-2 shows an example of a region tiled by these tiles.

As it turns out, there is no local-move property for these tiles. Figure 8-3 shows

Figure 8-1: The horizontal T-tetrominoes and the horizontal skew-tetrominoes.

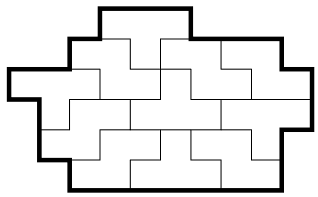


Figure 8-2: A tiling with tiles from $\mathcal{T}$.

a region for which local connectivity fails. There are actually five ways to tile this region, but four of them are very similar to one another—they result from flipping the top two or bottom two tiles in the tiling on the right of Figure 8-3. The tiling on the left, however, is not near any other tiling. Hence any set of local moves would have to include a single move which changes the tiling on the left to a tiling like that on the right. It is clear how this example can be generalized to an arbitrarily size, thus no finite set of local moves is sufficient.

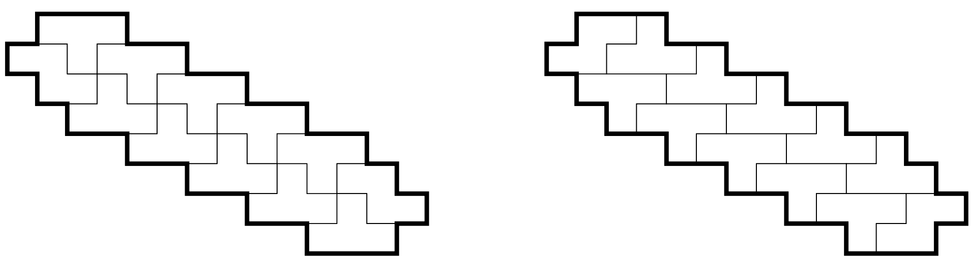


Figure 8-3: Local connectivity cannot hold for this region.

8.2.1 Tile invariants

Fix a region Γ, and let τ be a tiling of Γ. Recall that for $1 \leq i \leq 4$, $a_i(\tau)$ denotes the number of occurrences of tile t_i in the tiling. In this case, t_1 and t_2 are the

T-tetrominoes and t_3 and t_4 are the skew-tetrominoes, as shown in Figure 8-1. Let

$$
\begin{aligned}
b_1(\tau) &= a_1(\tau) + a_2(\tau) + a_3(\tau) + a_4(\tau) \\
b_2(\tau) &= a_1(\tau) - a_2(\tau) \\
b_3(\tau) &= a_1(\tau) \\
b_4(\tau) &= a_4(\tau)
\end{aligned}
$$

Theorem 8.1 *We have that b_1 is invariant (mod ∞) and b_2 is invariant mod 4.*

Proof: The first assertion is easy; b_1 always equals the area of Γ divided by 4. As for the second assertion, number the cells of the grid according to the following pattern (all taken mod 8).

$$
\begin{array}{ccccc}
1 & 1 & 1 & 1 & 1 \\
3 & 3 & 3 & 3 & 3 \\
5 & 5 & 5 & 5 & 5 \\
7 & 7 & 7 & 7 & 7 \\
1 & 1 & 1 & 1 & 1 \\
3 & 3 & 3 & 3 & 3
\end{array}
$$

One can verify that any placement of a t_1 tile will cover cells summing to 2 mod 8, while any t_2 tile will cover cells summing to -2 mod 8. Any skew tetromino (t_3 or t_4) will cover cells summing to 0 mod 8. Hence the sum of all the cells in Γ will equal $2 \cdot b_2$ (mod 8), thus b_2 is invariant modulo 4. ■

Theorem 8.2 *The invariants in Theorem 8.1 determine the integral lattice $\Lambda_{\mathcal{T},\mathcal{R}}$ completely for $\mathcal{R}_{all}$.*

Proof: First we will give an example which shows that b_4 can vary freely (Figure 8-4).

In this case, the first tiling is $(0, 1, 0, 1)$ in the a-basis, or $(2, -1, 0, 1)$ in the b-basis. The second tiling is $(0, 1, 1, 0)$ in the a-basis, or $(2, -1, 0, 0)$ in the b-basis. So b_4 can vary freely.

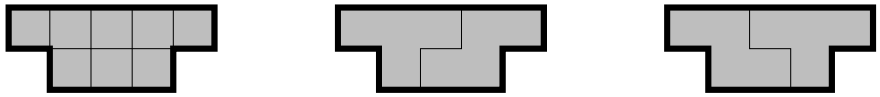


Figure 8-4: b_4 can vary freely.

For b_3, consider the region in Figure 8-5. The first tiling is $(0,0,2,0)$ in the a-basis, and $(2,0,0,0)$ in the b-basis. The second tiling is $(1,1,0,0)$ in the a-basis, and $(2,0,1,0)$ in the b-basis. Thus b_3 can vary freely.

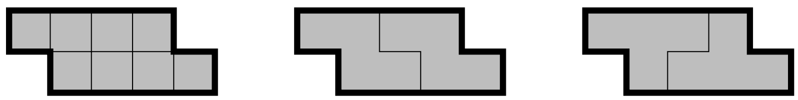


Figure 8-5: b_3 can vary freely.

As for b_2, consider the region shown in Figure 8-6. The first tiling is $(0,2,1,2)$ in the a-basis, and is $(5,-2,0,2)$ in the b-basis. The second tiling is $(2,0,1,2)$ in the a-basis, and is $(5,2,2,2)$ in the b-basis. We already saw that b_3 can vary freely, thus b_2 can vary by exactly 4. ■

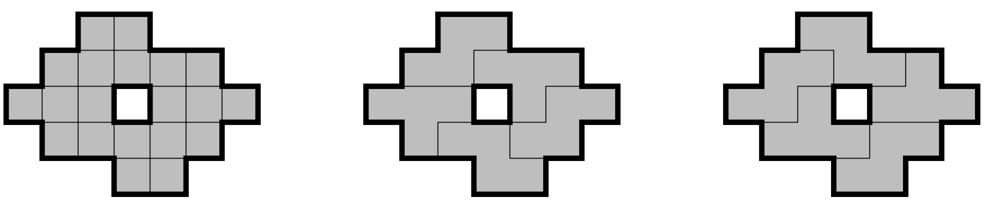


Figure 8-6: b_2 can vary by 4.

This defines all the tile invariants for this set of tiles. However, if we restrict our attention to simply-connected regions, then a new tile invariant appears. Notice that in the preceding proof, we made use of a non-simply connected region (Figure 8-6) as an example of a region where b_2 could vary by exactly 4. It turns out that this cannot happen if Γ is simply connected.

Theorem 8.3 *If Γ is a simply-connected region, then b_2 is invariant mod ∞.*

We will prove this using a boundary-word argument. Adorn the edges of the unit grid with labels and directions as follows. Label every horizontal edge with an a, and label every vertical edge with a b. Direct every horizontal edge to point east. Let all vertical edges with an even x-coordinate point north, and let all vertical edges with an odd x-coordinate point south (first panel of Figure 8-7). Now define the boundary word $w(\Gamma)$ as follows. Walk along the boundary of Γ in a counterclockwise direction. When you walk along an edge labelled a in the direction of that edge, write down the letter a. When you walk along an edge labelled a in the reverse direction, write down the letter a^{-1}. Do a similar thing for edges labelled b. Consider $w(\Gamma)$ to be an element of the free group on generators a and b. For example, take the region Γ shown in the middle panel of Figure 8-7. If we begin at the southwest corner, the boundary word will be

$$w(\Gamma) = aaab^{-1}aba^{-1}b^{-1}a^{-1}a^{-1}bba^{-1}b^{-1}.$$

From the boundary word $w(\Gamma)$ we can construct a 2-dimensional path $\pi(\Gamma)$ as follows. When you encounter the letter a, go east. When you encounter a^{-1}, go west. When you encounter b, go north. And when you encounter b^{-1}, go south. The third panel of Figure 8-7 shows the path $\pi(\Gamma)$ for our example region.

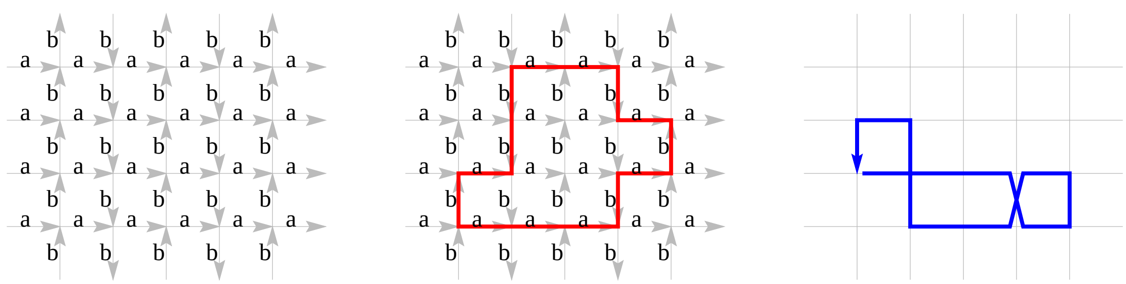


Figure 8-7: The labelling of the edges of the grid, a region Γ on the grid, and the path $\pi(\Gamma)$.

Lemma 8.4 *Let t be a tile in $\mathcal{T}$ located anywhere on the unit grid. Then $\pi(t)$ forms a closed loop with signed area 0.*

Proof: There are essentially eight different things to check here, because there are four different types of tile, and each may be situated at an odd or an even x-coordinate.

Each of these cases, and their respective paths, are shown in Figure 8-8 (we will refer to them again). Each is a closed loop with signed area 0, proving the lemma. ■

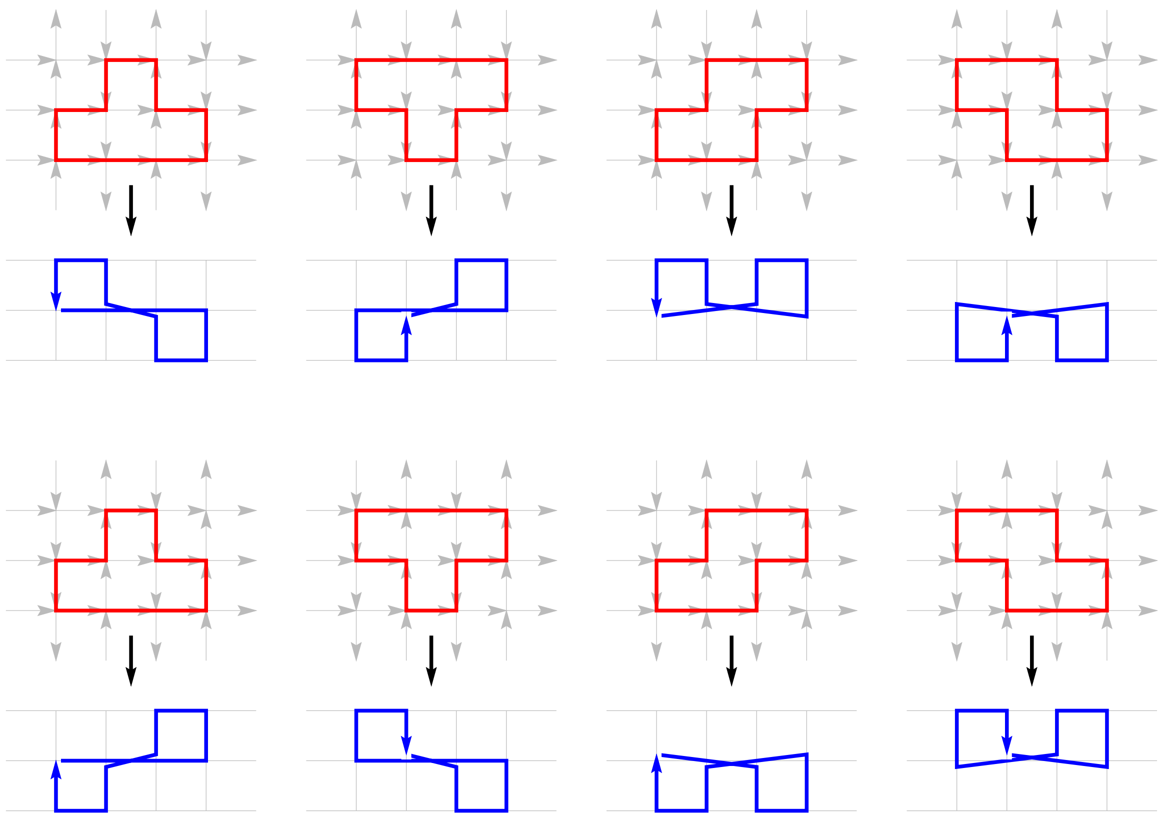


Figure 8-8: The eight different possible tile placements t (red), and the closed paths $\pi(t)$ they correspond to (blue).

Corollary 8.5 *If Γ is a simply connected region tileable by $\mathcal{T}$, then $\pi(\Gamma)$ forms a closed loop with signed area 0.*

Let us now assign a weight to each cell of the unit grid. In this case, we will assign each cell a weight which is equal to the y-coordinate of its bottom edge. For a closed loop, we define its *weighted area* to be the sum of the weights of those cells encircled counterclockwise, minus the sum of the weights of those cells encircled clockwise. (More precisely, the weighted area is $\sum_c \iota(c)\omega(c)$, where the sum is over all cells c, and $\iota(c)$ is the weight of c, and $\omega(c)$ is the path's winding number around c.)

For a simply connected tileable region Γ, let $\zeta(\Gamma)$ denote the weighted area of $\pi(\Gamma)$. Notice that since $\pi(\Gamma)$ has signed area 0, its weighted area will not change if the loop

$\pi(\Gamma)$ is translated in the plane. It is also interesting to notice that if Γ is translated in the plane (a translation of 1 in the x-direction is the only case that matters), then $\zeta(\Gamma)$ still remains constant. (In this case, $\pi(\Gamma)$ is reflected vertically, so high-weight cells switch with low-weight cells, but clockwise becomes counterclockwise and vice versa, leaving the weighted area unchanged.)

Lemma 8.6 *If t is an upward-pointing T-tetromino, $\zeta(t) = 1$. If t is an downward-pointing T-tetromino, $\zeta(t) = -1$. If t is a horizontal skew-tetromino, $\zeta(t) = 0$.*

Proof: The proof is clear from inspection of Figure 8-8. ∎

From this, the proof of Theorem 8.3 is immediate. It is clear from Lemma 8.6 and from the additive nature of ζ that b_2 is precisely $\zeta(\Gamma)$, which does not depend on τ, hence b_2 is invariant mod ∞.

8.3 Horizontal T-tetrominoes and horizontal dominoes

Now let us consider another set of tiles which looks somewhat similar, but has a much different character. Consider the set of tiles which consists of the horizontal domino, and the two types of horizontal T-tetrominoes. We will define a height function for these tiles, and use it to prove that local connectivity holds for tilings of simply connected regions with these tiles.

Let $\mathcal{T}$ denote the set of tiles consisting of the horizontal domino, the upward-pointing T-tetromino, and the downward-pointing T-tetromino, as shown in Figure 8-9. Figure 8-10 shows a tiling of a region with these tiles. Let $\mathcal{L}$ be the set of local moves depicted in Figure 8-11.

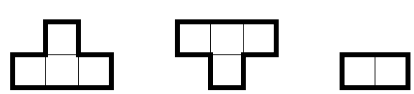


Figure 8-9: The horizontal T-tetrominoes, and the horizontal domino.

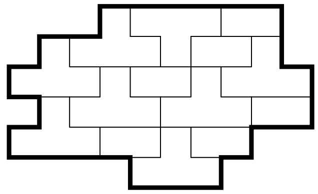


Figure 8-10: A tiling of a region with these tiles.

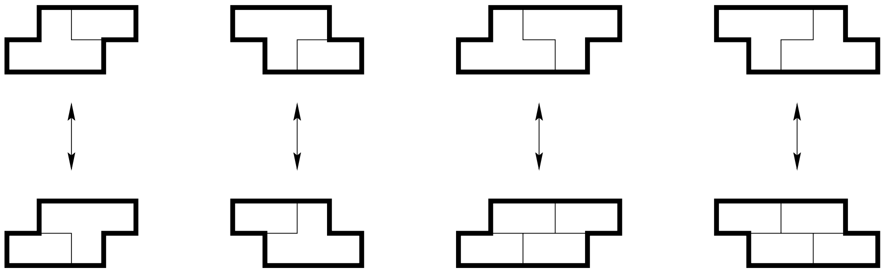


Figure 8-11: Our set of local moves.

Theorem 8.7 *Tilings of simply-connected regions with the tile set $\mathcal{T}$ have local connectivity with respect to $\mathcal{L}$.*

It should be noted that this theorem fails if the region is not simply connected. Figure 8-12 shows an arbitrarily large region which has only two tilings, thus showing that no local-move property can possibly hold for non-simply connected regions with this tile set.

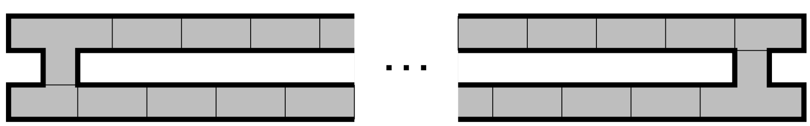


Figure 8-12: A non-simply connected region which can be tiled in only two ways.

8.3.1 The height function

For this set of tiles we define a height function which is analogous to the classical height function for domino tilings. As in the case of dominoes, our height function assigns an integer value to each vertex. This height function has the following properties:

- It is defined only for tilings of simply connected regions.

- It is unique up to an additive constant.
- If Γ is a tileable region, then the heights of the points on the boundary of Γ do not depend of the particular tiling of Γ.

Let Γ be a simply connected tileable region, and let τ be a tiling of this region. As in the case of dominoes, we use a local rule to define the height function corresponding to τ. Begin by assigning an arbitrary height to a point on the boundary. Now move along edges of tiles to assign heights to the remaining points, by the following rules:

- As one moves east or west along an edge, the height remains the same.
- As one moves north (resp. south) along an edge with an even x-coordinate, the height increases (resp. decreases) by 1.
- As one moves north (resp. south) along an edge with an odd x-coordinate, the height decreases (resp. increases) by 1.

Collectively, we will refer to these as *condition* $\star$.

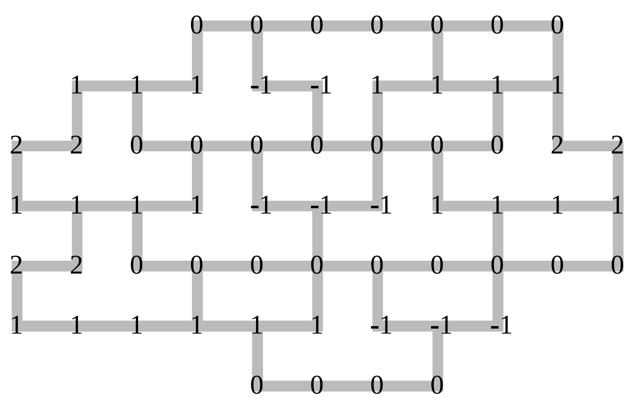


Figure 8-13: Heights for points in our sample tiling.

Lemma 8.8 *This height function is well-defined for tilings of simply connected regions.*

Proof: It suffices to show that upon completing any closed loop, the net change in height is 0. Since Γ is simply connected, any closed loop can be written as a sum of loops around tiles. Hence it suffices to show that the net change in height is 0 for loops which are the boundary of a single tile. This is easily seen to be true. ■

Define two points (x_1, y_1) and (x_2, y_2) to be a *visible pair* if the straight line between the points lies entirely within the region Γ (including its boundary).

Lemma 8.9 *Let Γ be a simply connected, tileable region. A function h on the lattice points of Γ is a height function if and only if the following hold.*

1. *Values of h on the boundary of Γ obey condition $\star$.*

2. *Visible pairs of points which are vertically or diagonally (but not horizontally) adjacent have values differing by exactly 1.*

3. $h(2i, j) \leq h(2i + 2, j + 1) + 1$ *whenever* $(2i, j)$ *and* $(2i + 2, j + 1)$ *form a visible pair.*

4. $h(2i, j) \leq h(2i - 2, j + 1) + 1.$ *whenever* $(2i, j)$ *and* $(2i - 2, j + 1)$ *form a visible pair.*

5. $h(2i - 1, j) \geq h(2i + 1, j + 1) - 1.$ *whenever* $(2i - 1, j)$ *and* $(2i + 1, j + 1)$ *form a visible pair.*

6. $h(2i - 1, j) \geq h(2i - 3, j + 1) - 1.$ *whenever* $(2i - 1, j)$ *and* $(2i - 3, j + 1)$ *form a visible pair.*

7. $h(i, j) = h(i + 1, j)$ *or* $h(i, j) = h(i - 1, j).$ *whenever* $(i + 1, j)$ *and* $(i - 1, j)$ *form a visible pair.*

We will refer to these as *conditions 1-7.* Before going any further, let us notice a few things about these conditions. First, notice that condition 2 implies that all points with an even y-coordinate will have even heights, and all points with an odd y-coordinate will have odd heights (or vice versa). Second, notice that there are certain symmetries among the conditions 3 through 6. Namely, if Γ is a region, and h is a function which satisfies conditions 1-7, then upon reflecting the entire picture across the y-axis, h will still obey conditions 1-7. Also, if we instead reflect Γ across the x-axis and shift the picture 1 unit to the right, the resulting function h will still obey conditions 1-7 as well.

Now let us prove this lemma.

Proof: First let us prove that these relations hold when h is a height function. Assume that we have some tiling of the region Γ, and h is the corresponding height function.

The first condition trivially holds. As for the second condition, consider two points p and q which are diagonally adjacent. Since they form a visible pair, the cell between them must belong to some tile. There are essentially ten different configurations that this tile can have with respect to p and q (i.e., the cell between p and q can be the leftmost cell of a domino, or the rightmost cell of a domino, or the leftmost cell of an upward-pointing T-tetromino, etc.). It is straightforward to verify that in each of these cases, the heights for p and q will differ by exactly 1. Similarly, if p and q are vertically adjacent, then again either the cell to the left of them or the cell to the right of them must belong to some tile. There are a few cases to check, but in each instance the heights differ by exactly 1.

Now consider condition 3. Let us suppose we had points which violated this condition. Let us call them p and q. By condition 2, we know that the heights of p and q must differ by 3, and that the heights of the neighboring points must be as indicated in Figure 8-14 (we know all these points lie in Γ since p and q are a visible pair). The dotted gray lines in the figure cannot be tile boundaries, since the heights of their endpoints don't conform to condition $\star$. But this is clearly impossible, given the shapes of the allowed tiles.

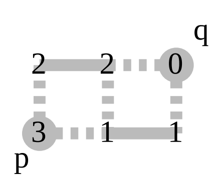


Figure 8-14: Points p and q violate condition 3.

Conditions 4, 5, and 6 are essentially the same as the above, just reflected and perhaps shifted by 1 in the x direction.

As for the final condition, suppose $h(i,j) \neq h(i+1,j)$. Then the segment between these points cannot be a tile boundary, hence it must be the segment across the middle of a T-tetromino (Figure 8-15). There are two such configurations, and in either case, $h(i,j) = h(i-1,j)$.

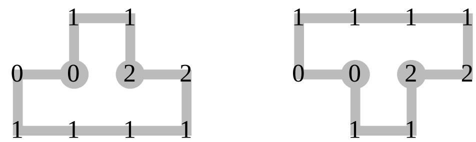


Figure 8-15: The case when $h(i,j) \neq h(i+1,j)$

The proves the "only if" direction.

For the "if" direction, let h be a function which satisfies the conditions in the lemma. Our plan will be to construct a tiling from this function. This is done by drawing a tile boundary everywhere the heights agree with condition $\star$. (In other words, draw an edge between (i,j) and $(i+1,j)$ if they have the same height, draw an edge between $(2i,j)$ and $(2i,j+1)$ if the height of the latter point is 1 more the the height of the former point, etc.) We need to show that the tiling constructed in this way is a valid tiling of Γ.

Let c be any cell in Γ. Assume without loss of generality that the left side of c has an even x-coordinate, and that the lower-left corner of c has coordinates $(2i,j)$ and has height 0. There are now six possibilities for the heights of points around c, which are shown in Figure 8-16. Let us label these cases A through F respectively.

Figure 8-16: Cases A through F for the heights of points near cell c.

Let us consider case B first. The cell directly above c must lie in the region Γ, and we must have $h(2i,j+2) = h(2i+1,j+2) = 0$ by condition 2 (see Figure 8-17). Then the cells northeast and northwest of c must also lie in Γ, and we must have $h(2i-1,j+1) = 1$ and $h(2i+2,j+1) = -1$, by condition 7. Furthermore, we must have $h(2i+2,j+2) = 0$ by condition 3, and $h(2i-1,j+2) = 0$ by condition 6. So in this case, the edge boundaries form a downward-pointing T-tetromino whose bottom cell is c. Case E is the same thing, just upside down.

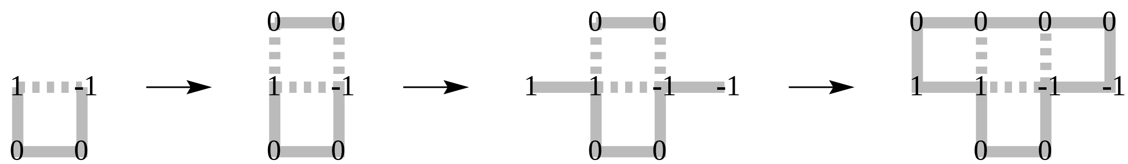


Figure 8-17: Deducing nearby heights in case B.

Now let us tackle case C (Figure 8-18). By condition 2 we get $h(2i, j+2) = h(2i+1, j+2) = 0$. By condition 7, we get $h(2i-1, j+1) = -1$ and $h(2i+2, j+1) = 1$. And we get $h(2i-1, j) = 0$ by condition 5, and $h(2i+2, j) = 0$ by condition 4. So in this case, the edge boundaries form an upward-pointing T-tetromino whose middle cell is c. Case F is the same thing, just upside down.

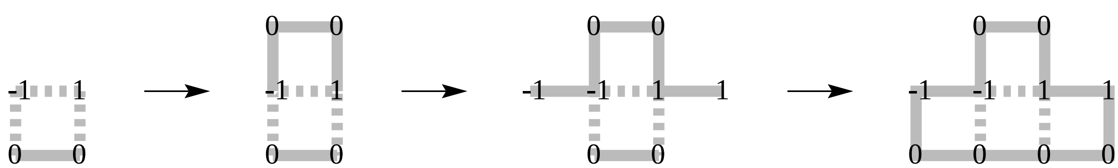


Figure 8-18: Deducing nearby heights in case C.

For case A we have a few subcases (Figure 8-19), which we will call subcases A1, A2, and A3. In case A1, we get $h(2i+3, j) = 0$ by condition 6, we get $h(2i+3, j+1) = -1$ by condition 7, and $h(2i+1, j+2) = h(2i+2, j+2) = 0$ by condition 2. So we get an upward-pointing T-tetromino whose leftmost cell is c. In case A2, we get the same thing, just upside-down. And in case A3, we are already done, since the edge boundaries form a domino whose leftmost cell is c.

Case D follows in the same way.

We have shown that if h is any function which satisfies the 7 conditions in the lemma, then drawing the appropriate edge boundaries only creates regions which are horizontal T-tetrominoes or horizontal dominoes, thus it produces a valid tiling of Γ. Hence h is in fact a height function. ∎

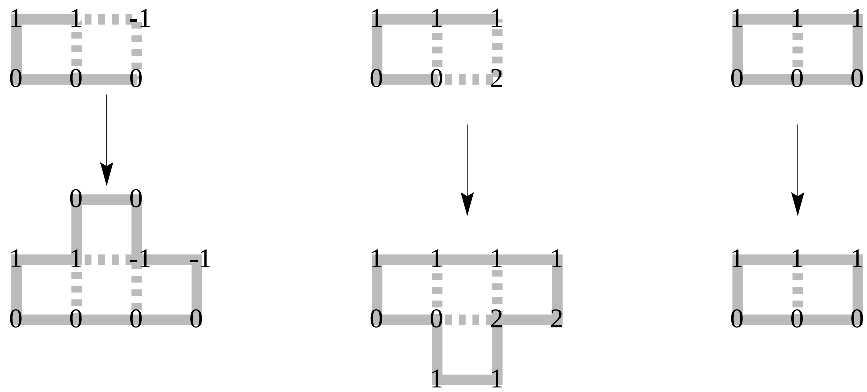


Figure 8-19: Cases A1, A2, and A3.

8.3.2 The lattice of height functions

Fix a simply-connected region Γ. Let h_1 and h_2 be height functions on Γ. Define a partal ordering $\leq$ by the rule that $h_1 \leq h_2$ iff $h_1(x) \leq h_2(x)$ for all x in Γ.

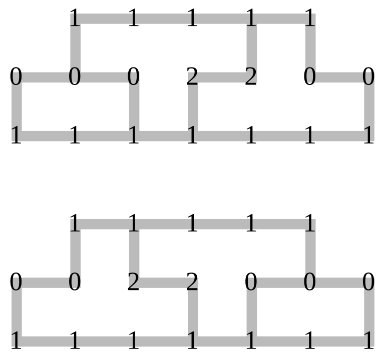
Theorem 8.10 *The set of height functions on Γ, with this partial ordering, is a lattice.*

Proof: In order to prove this, we need to show the existence of a unique greatest lower bound $h_1 \wedge h_2$ for any two height functions h_1 and h_2. (The case of the least upper bound will be identical.)

Let h_1 and h_2 be any two height functions on Γ. Define the function h_{min} by the rule $h_{min}(x) = \min\{h_1(x), h_2(x)\}$. If h_{min} is a valid height function, then we are done, since h_{min} is obviously a lower bound, and every other lower bound must be less than it. Unfortunately, h_{min} will not be a height function in general.

Let us see how h_{min} can fail to be a height function. Recall conditions 1-7 in Lemma 8.9. Observe that h_{min} satisfies the first six of them, because h_1 and h_2 satisfy them. However, it is possible that h_{min} violates condition 7 (Figure 8-20).

Let us call a point (i, j) an *isolated peak* of h_{min} if $h_{min}(i-1, j) = h_{min}(i, j) - 2 = h_{min}(i + 1, j)$. Let us call a point (i, j) an *isolated valley* of h_{min} if $h_{min}(i - 1, j) = h_{min}(i, j) + 2 = h_{min}(i+1, j)$. Notice that it is impossible for h_{min} to have any isolated valleys, because this would force either h_1 or h_2 to have an isolated valley. So we only need to concern ourselves with isolated peaks.



	1	1	1	1	1	
0	0	0	2	0	0	0
1	1	1	1	1	1	1

Figure 8-20: Height functions h_1 and h_2 (on left), and h_{min}, which is not a height function.

Let us define h^* to be the same as h_{min}, but with the isolated peaks smoothed over. More precisely, define $h^*(p) = h_{min}(p) - 2$ whenever p is an isolated peak, and define $h^*(p) = h_{min}(p)$ for all other points p. This function h^* will turn out to be a height function, and will be the desired greatest lower bound.

First we must show that h^* is a valid height function. Recall that h_{min} satisfied conditions 1-6, so points that were unchanged between h_{min} and h^* will still satisfy conditions 1-6. And we specifically designed h^* to satisfy condition 7. So all we need to check is that conditions 1-6 were not violated by decreasing the height of isolated peaks.

Let p be an isolated peak of h_{min}. Assume without loss of generality that $h_{min}(p) = 2$. Then the values of h_{min} for points around p (if they lie in Γ) must be as shown in Figure 8-21. The points marked with a $+$ may have values of either 1 or -1. The point p cannot lie on the boundary of Γ, so condition 1 will not be violated. And by examining Figure 8-21, we can see that conditions 2-6 will not be violated by setting $h^*(p)$ to 0. Hence h^* will be a valid height function.

+	1	1	1	+
0	0	2	0	0
+	1	1	1	+

Figure 8-21: The values of h_{min} for the points near p.

All that remains is to show that h^* is the greatest lower bound for h_1 and h_2. Clearly it is a lower bound, since it is less than h_{min}. Suppose h_{other} is a different lower bound, and $h_{other} \not\leq h^*$. Certainly $h_{other} \leq h_{min}$, so the only points at which h_{other} could exceed h^* would be those points which are isolated peaks of h_{min}. But then these points would also have to be isolated peaks of h_{other}, violating the fact that h_{other} is a height function.

So $h^* = h_1 \wedge h_2$, and the height functions on Γ form a lattice. ■

Unlike the case of dominoes or T-tetrominoes, this lattice is not distributive. Figure 8-22 shows an example of height functions f, g, and h such that $(f \vee g) \wedge h \neq (f \wedge h) \vee (g \wedge h)$.

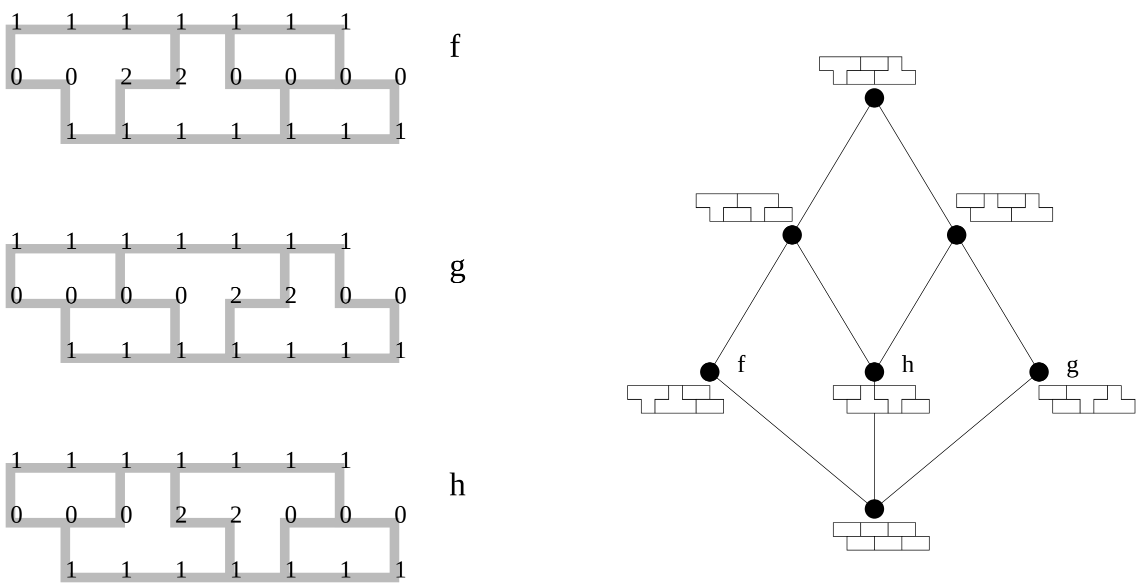


Figure 8-22: Height functions f, g, and h (left) provide a counterexample to the distributive laws. On the right is the Hasse diagram for the nondistributive lattice formed by the height functions of this region.

8.3.3 Local connectivity

Let us use this height function to prove Theorem 8.7. First we will examine the effect of a local move on a height function.

Lemma 8.11 *Let Γ be a simply-connected region, and let h_1 and h_2 be two height functions on Γ. If p is a point of Γ, and $h_1(x) = h_2(x)$ for all $x \neq p$, then the* *tilings corresponding to h_1 and h_2 differ by a local move. If p_1 and p_2 are horizontally adjacent points of Γ, and $h_1(x) = h_2(x)$ for all x other than p_1 or p_2, then the tilings corresponding to h_1 and h_2 differ by one or two local moves.*

Proof: Let us consider the first case, where h_1 and h_2 differ only at a single point p. Say this point has coordinates (i, j), and without loss of generality, assume i is even. Suppose $h_1(p) = U - 1$ and $h_2(p) = U + 1$, for some U. Of the points $(i+1, j)$ and $(i-1, j)$, one must have height $U - 1$ and the other must have height $U + 1$, or else condition 7 would be violated for either h_1 or h_2. Then applying conditions 1-7 for height functions, the points near p must have heights as indicated in Figure 8-23. Thus we can see that the corresponding tilings differ by a local move.

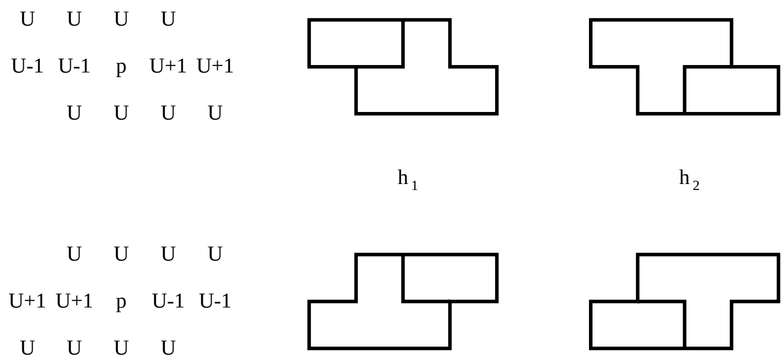


Figure 8-23: Height functions which differ at a single point.

Now for the second case, assume that h_1 and h_2 differ at a pair of adjacent points, p_1 and p_2. Let p_1 have coordinates (i, j), and let p_2 have coordinates $(i + 1, j)$, and assume without loss of generality that i is even. By looking at the possible cases, one can verify that the only way this is possible is if $h_1(p_1) = h_1(p_2) = U - 1$ and $h_2(p_1) = h_2(p_2) = U + 1$ for some U (or vice versa). Let us consider the following two subcases.

In the first subcase, we assume $(i-1, j)$ and $(i+2, j)$ have the same height. Using rules 1-7 for height functions, we can deduce the heights of other points near these. We must have one of the situations depicted in Figure 8-24. In either case, the tilings corresponding to h_1 and h_2 differ by a local move.

In the second subcase, we take $(i-1, j)$ and $(i+2, j)$ to have different heights. Using rules 1-7 for height functions, we can deduce the heights of other points near these. We must have one of the situations depicted in Figure 8-25. This time, the tilings corresponding to these height functions do not differ by a local move, but it is easy to see that they are connected by a sequence of two local moves. ■

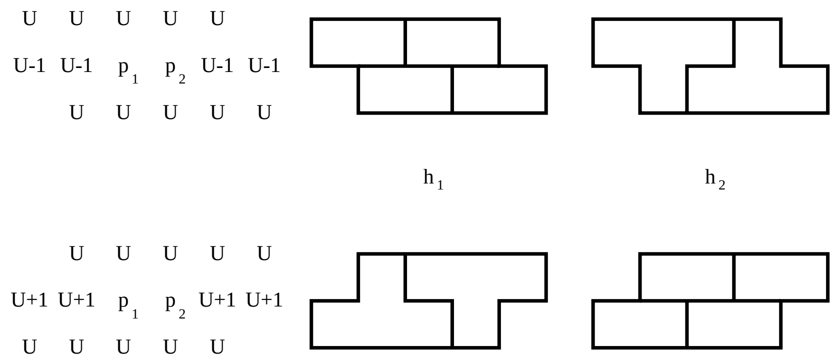


Figure 8-24: Height functions which differ at points p_1 and p_2 (first subcase).

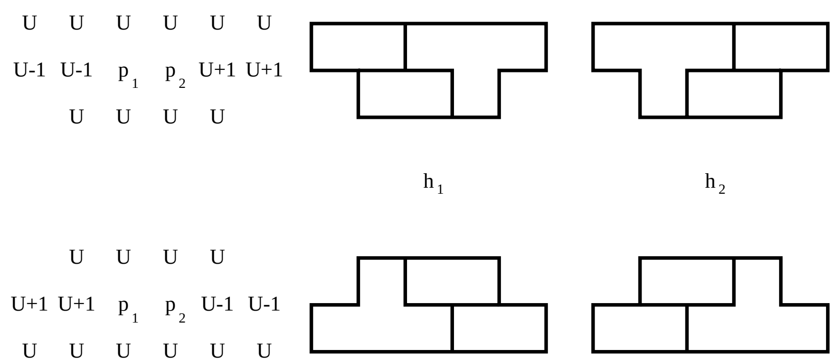


Figure 8-25: Height functions which differ at points p_1 and p_2 (second subcase).

Now that we understand the effect of local moves on height functions, let us use these height functions to prove Theorem 8.7. For convenience in what follows, we may refer to height functions as though they were tilings, and talk about performing local moves on a height function.

Proof: We know that the set of all height functions on Γ forms a lattice, hence there must be a unique lowest height function h_0. Let h be any height function other than h_0. We will show that there is always a local move we can apply to h which will yield a lower height function h^*. By repeatedly performing such moves, we will be able to convert any height function to h_0, hence all height functions (or tilings) will be connected by local moves.

Again, let h be any height function other than h_0. Let us color blue those points p of Γ where $h(p) = h_0(p)$, and color red the points where $h(p) > h_0(p)$. Notice that all points on the boundary of Γ are blue. Since $h \neq h_0$, there is at least one red point. Let T be the height of the highest red point.

Define a *red ridge* to be a maximal collection of red points with height T lying consecutively along a horizontal line. Let Υ be any red ridge. We will show that we can find either one or two points of Υ such that reducing their heights from T to $T - 2$ results in a valid height function h^*. By Lemma 8.11, we know that h and h^* must be connected by a local move (or two). This will prove the theorem.

Suppose we decrease the heights of some points belonging to Υ by 2, forming h^*. Certainly h^* will not violate condition 1, since no red points are on the boundary of Γ. Suppose some point $x \in \Upsilon$ whose height we changed now violates some condition 2-6. This could only happen if there were a point r in the row above or below Υ with $h(r) = T + 1$. Such a point would have to be blue, since the maximum red height which appears is T. Since it is blue, we must have $h_0(r) = T + 1$. But then we would have $h_0(x) \geq T$, by the same condition that we assumed was violated by h^*. This contradicts the fact that x is red. Thus the only condition that h^* could possibly violate is condition 7. Hence we may safely ignore all but the row containing Υ.

Let $x_1, x_2, \ldots, x_k$ be the points of Υ from left to right, let p be the point left of x_1, and let q be the point right of x_k. Then let p' be the point left of p, and let q' be the point right of q, if they belong to Γ. (So from left to right, we have $p', p, x_1, x_2, \ldots, x_k, q, q'$.) Notice that p and q cannot be red $T + 2$'s, since T is the largest red height that appears. They cannot be red T's, or else Υ would not be maximal. Observe also that p and q cannot be blue $T + 2$'s. If p were a blue point of height $T + 2$, then we would have $h_0(p) = T + 2$, and thus we would have $h_0(x_1) \geq T$. But x_1 is a red T, which is a contradiction. So p and q can be red or blue points with height $T - 2$, or blue points with height T.

Case 1: Suppose $k = 1$. Then we cannot have $h(p) = h(q) = T - 2$, or else h would violate condition 7. Suppose we had $h(p) = h(q) = T$. Then p and q must both be blue, so $h_0(p) = h_0(q) = T$. But x_1 is red, so $h_0(x_1) < T$. This violates condition 7. So we must have $h(p) = T - 2$ and $h(q) = T$, or vice versa (which will be the same by symmetry). Then q is blue, so we must have $h_0(x_1) = T - 2$. Thus by condition 7, $h_0(q') = T$. So $h(q') \geq T$, and since T is the largest red height that appears, we must have $h(q') = T$. Now we form h^* by reducing the height of x_1 to $T - 2$. This does not create an isolated valley at x_1, and it does not leave an isolated peak at q. So h^* is a valid height function.

Case 2: Suppose $k \geq 2$ and $h(p) = h(q) = T$. Then p and q must both be blue. So $h_0(p) = h_0(q) = T$, and we must have $h_0(x_1) = h_0(x_k) = T - 2$. Then in order for h_0 not to violate condition 7, we must have $h_0(p') = h_0(q') = T$. Thus $h(p') = h(q') = T$. Now form h^* by reducing the heights of x_1 and x_2 to $T - 2$. This does not create any isolated peaks or valleys, so h^* is a valid height function.

Case 3: Suppose $k = 2$ and $h(p) = h(q) = T - 2$. Form h^* by reducing the height of x_1 and x_2 to $T - 2$. This does not violate condition 7, so h^* is a valid height function.

Case 4: Suppose the situation is not one of those in cases 1-3. Either $h(p)$ or $h(q)$ must equal $T - 2$, or both. Assume without loss of generality that $h(p) = T - 2$. Let s be the point to the right of x_2 (so $s = x_3$ if $k \geq 2$, and $s = q$ if $k = 2$). Then $h(s) = T$ (because if $s = q$ and $h(s) = T - 2$, then we would be in Case 3). Now form h^* by reducing the height of x_1 to $T - 2$. This does not create an isolated peak or valley, hence h^* is a valid height function.

We have shown that for any height function $h \neq h_0$, there is another lower height function which differs from h in at most two places. Hence we can always apply local moves to decrease the height. Thus all height functions can be reduced to h_0, thus all height functions are connected by local moves. ∎

8.3.4 Conclusion

It is entirely possible that one could construct a shorter proof of Theorem 8.7 by an *ad hoc* argument. However, by doing so, one would likely miss out on the lattice structure and other interesting features of this approach. The interesting thing about this result is not just that a certain set of tiles has a local-move property, but that the notion of a height function can be made to work even in situations where the definitions are not as simple as those for the domino height function.

In some sense, condition 7 is the most surprising thing about this height function. Condition 1 just ensures that the values on the boundary will be constant, and conditions 2-6 simply dictate that heights of points which are near one another should not vary too much. But condition 7 is something which is harder to explain, since it has no analogue in the domino height function scheme. It destroys the distributivity of the lattice, and makes us do a little bit of extra work just to define meets and joins. Yet despite this, we are still able to carry the approach through to the end and prove local connectivity.

It would be very interesting to see other sets of tiles for which one could define a height function like this one. Was it just good luck that our condition 7 did not destroy the entire approach? Or was it just bad luck that we needed such a poorly-behaved condition at all?

Chapter 9

Tiling with skew-tetrominoes

9.1 Introduction

Let $\mathcal{T}$ denote the set of *skew-tetrominoes*, as shown in Figure 9-1.

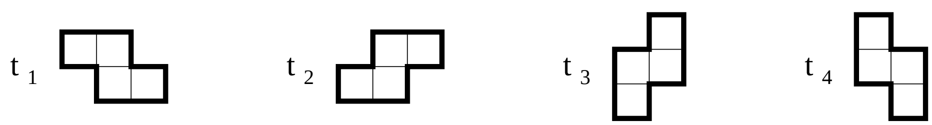


Figure 9-1: Skew-tetrominoes.

These tiles have been studied in [23]. The focus in that paper was using the notion of boundary words to prove the impossibility of tiling certain regions.

9.2 Local moves

Theorem 9.1 *The set of skew-tetrominoes does not have a local-move property for tilings of simply connected regions.*

Proof: Consider the region shown in Figure 9-2. It can be tiled in exactly six ways. In each of the two tilings shown, there is a small yellow area where local moves can be applied, but these are the only ones possible. In order for the skew tetrominoes to have a local move property, this entire region must be considered to be one local move. It is easy to see how this region can be generalized to a region of arbitrary size (while still admitting only six tilings), hence no finite set of local moves can suffice. ■

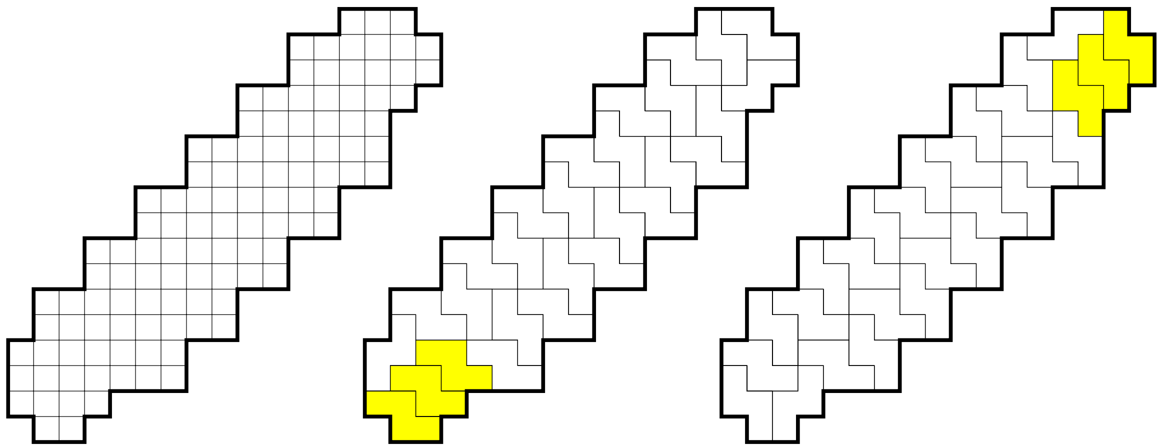


Figure 9-2: Two tilings which are not connected by local moves.

9.3 Tile invariants

Let Γ be a region, and let τ be a tiling of this region. As we did in chapter 3, we will let $a_i(\tau)$ denote the number of occurrences of the tile t_i in the tiling τ. (Figure 9-1 shows which tile is which.)

Define the b-basis as follows.

$$\begin{aligned} b_1(\tau) &= a_1(\tau) + a_2(\tau) + a_3(\tau) + a_4(\tau) \\ b_2(\tau) &= a_2(\tau) \\ b_3(\tau) &= a_3(\tau) \\ b_4(\tau) &= a_4(\tau) \end{aligned}$$

Theorem 9.2 *We have that b_1 is constant mod ∞. Also, b_2 is invariant mod 2, b_3 is invariant mod 2, and b_4 is invariant mod 2.*

This also implies that a_1 is invariant mod 2, by subtracting the last three invariants from the first one.

Proof: The first assertion is trivial; it is the area of Γ divided by 4.

For b_2, consider the following coloring argument. Assign values to the cells of the grid as follows:

0	1	2	3	4	5	6	7	0	1	2	3	4	5
5	4	3	2	1	0	7	6	5	4	3	2	1	0
6	7	0	1	2	3	4	5	6	7	0	1	2	3
7	6	5	4	3	2	1	0	7	6	5	4	3	2
4	5	6	7	0	1	2	3	4	5	6	7	0	1
1	0	7	6	5	4	3	2	1	0	7	6	5	4
2	3	4	5	6	7	0	1	2	3	4	5	6	7
3	2	1	0	7	6	5	4	3	2	1	0	7	6
0	1	2	3	4	5	6	7	0	1	2	3	4	5
5	4	3	2	1	0	7	6	5	4	3	2	1	0

Observe that any placement of a t_2 tile covers values summing to 4 mod 8. However, any placement of a t_1-tile, t_3-tile, or t_4-tile covers values summing to 0 mod 8. Let d be the sum of the values of the cells in Γ. If τ is a tiling of Γ, then $d \equiv 4 \cdot a_2(\tau) \pmod 8$, hence $a_2(\tau)$ (or $b_2(\tau)$) will be invariant mod 2.

The coloring arguments for b_3 and b_4 are the same, just rotated and/or reflected.

■

Theorem 9.3 *The invariants in Theorem 9.2 determine the integral lattice $\Lambda_{\mathcal{T},\mathcal{R}}$ completely for $\mathcal{R}_{all}$.*

Proof: We need to exhibit regions where b_2, b_3, and b_4 can each vary by exactly 2. Let us start with b_2.

Consider Figure 9-3. The first tiling uses 2 copies of t_1, while the second tiling uses 2 copies of t_2. In the b-basis, the first tiling is $(2, 0, 0, 0)$, while the second tiling is $(2, 2, 0, 0)$. Thus b_2 may vary by exactly 2.

Now for b_4, consider the non-simply-connected region shown in Figure 9-4. The first tiling is $(2, 1, 1, 0)$ in the a-basis, which is $(4, 1, 1, 0)$ in the b-basis. The second

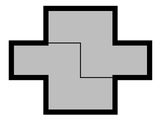


Figure 9-3: A region which shows that b_2 may vary by exactly 2.

tiling is $(0, 1, 1, 2)$ in the a-basis, which is $(4, 1, 1, 2)$ in the b-basis. Hence b_4 may vary by exactly 2.

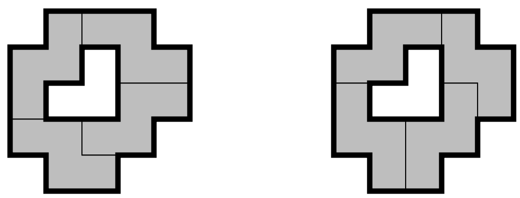


Figure 9-4: A region which shows that b_4 may vary by exactly 2, for non-simply-connected regions.

For b_3, consider the region shown in Figure 9-5. The first tiling is $(2, 0, 2, 0)$ in the b-basis, while the second tiling is $(2, 0, 0, 2)$ in the b-basis. Hence b_3 may vary by exactly 2. ■

Figure 9-5: A region which shows that b_3 may vary by exactly 2.

In order to prove that these invariants are the best possible, we had to resort to using a region which was not simply-connected. In fact, if we restrict our attention to just the simply-connected regions, we can improve upon these tile invariants slightly.

Theorem 9.4 *For simply connected regions, we have the additional relation that* $a_1(\tau) - a_2(\tau)$ *is invariant mod 4.*

Notice that this is equivalent to the assertion that $a_1(\tau) + a_2(\tau)$ is invariant mod 4. (We know that $a_2(\tau)$ is invariant mod 2, hence $2a_2(\tau)$ must be invariant mod 4. Adding this relation to the relation in Theorem 9.4 gives that $a_1(\tau) + a_2(\tau)$ is invariant mod 4.)

Let us set up a new basis to handle the case of simply-connected regions. Define

$$c_1(\tau) = a_1(\tau) + a_2(\tau) + a_3(\tau) + a_4(\tau)$$
$$c_2(\tau) = a_1(\tau) - a_2(\tau)$$
$$c_3(\tau) = a_2(\tau)$$
$$c_4(\tau) = a_4(\tau)$$

From Theorem 9.2, we know that c_1 is invariant mod ∞, and c_3 and c_4 are invariant mod 2. Incidentally, we also know that c_2 is invariant mod 2. We must show that for simply connected regions, c_2 is actually invariant mod 4.

Before proving Theorem 9.4, let us show that we cannot do any better than these invariants for simply-connected regions.

For c_4, take the tilings shown in Figure 9-5. The first tiling is $(2, 0, 0, 0)$ in the c-basis, while the second tiling is $(2, 0, 0, 2)$ in the c-basis. Thus c_4 may vary by exactly 2.

For c_3, consider the region shown in Figure 9-6. The first tiling is $(3, 2, 0, 1)$ in the a-basis, which is $(6, 1, 2, 1)$ in the c-basis. The second tiling is $(1, 0, 2, 3)$ in the a-basis, which is $(6, 1, 0, 3)$ in the c-basis. We already know that c_4 may vary by exactly 2, hence we have that c_3 may vary by exactly 2.

For c_2, take the tilings shown in Figure 9-3. The first tiling is $(2, 2, 0, 0)$ in the c-basis, while the second tiling is $(2, -2, 2, 0)$ in the c-basis. We know c_3 may vary by exactly 2, hence c_2 may vary by exactly 4.

Thus these invariants are the best possible.

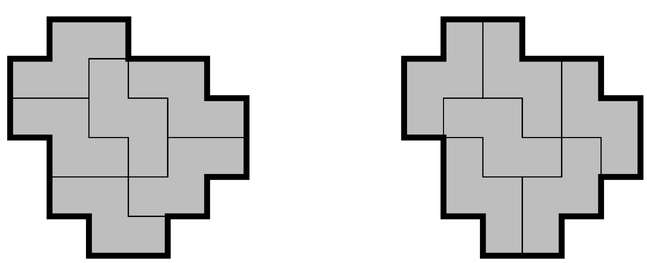


Figure 9-6: Two tilings of a simply-connected region.

Now let us prove Theorem 9.4.

Proof: Our approach will be to define boundary words for regions in the plane, then use these boundary words to create closed paths in a different way.

Let Γ be a region in the plane. Begin at any point on its boundary, and travel counterclockwise. Write the letter a (resp. a^{-1}) every time you move east (resp. west) along an edge having an even y-coordinate. Write the letter b (resp. b^{-1}) every time you move east (resp. west) along an edge having an odd y-coordinate. Don't write anything when you move north or south. Let us call the resulting word $w(\Gamma)$.

Now let us translate these words into alternate closed paths. We do this in a straightforward manner—we move east, west, north, or south whenever we see the letter a, a^{-1}, b, or b^{-1}, respectively. This draws out a path $\pi(\Gamma)$ in the alternate plane. Notice that if Γ is a tile, then $\pi(\Gamma)$ will be a closed path (see Figure 9-7). Hence if Γ is tileable, then $\pi(\Gamma)$ will be a closed path.

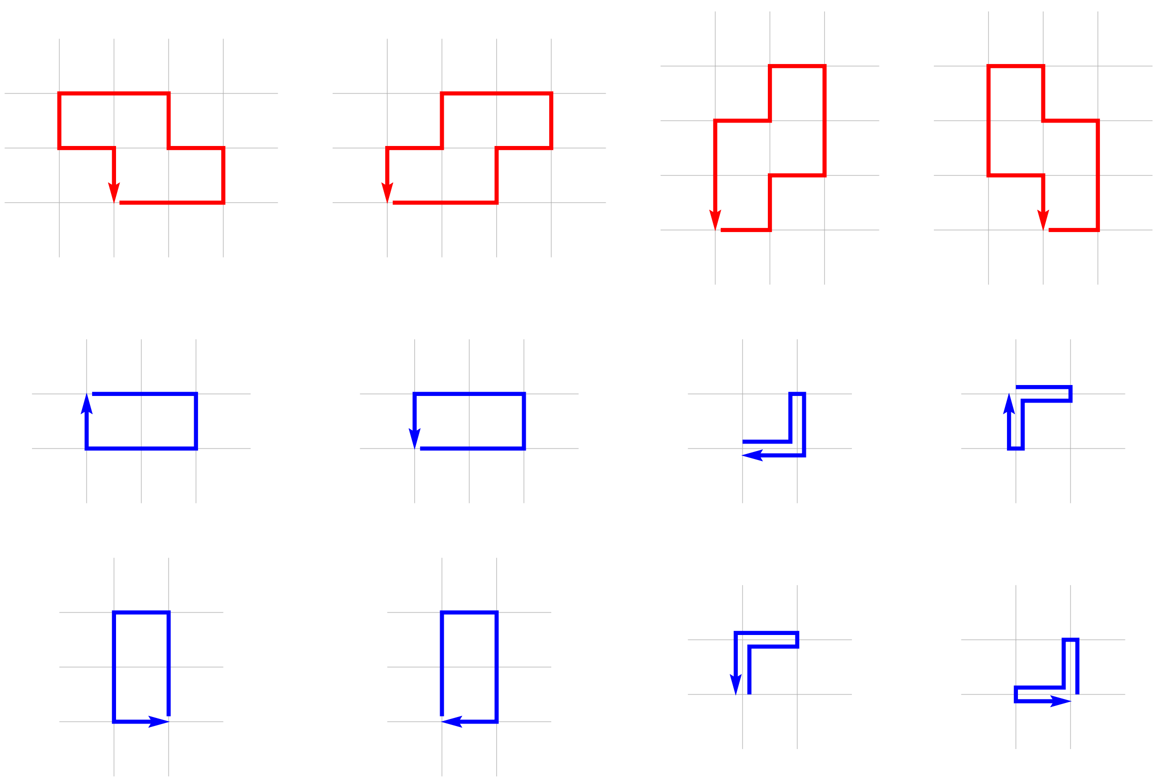


Figure 9-7: The tiles t, and the paths $\pi(t)$. The two paths for each tile correspond to whether the bottom of the original figure has an even or odd y-coordinate.

Now label cells in the new plane with values as shown below (all taken mod 4).

0	3	0	3	0	3	0	3
1	2	1	2	1	2	1	2
0	3	0	3	0	3	0	3
1	2	1	2	1	2	1	2
0	3	0	3	0	3	0	3
1	2	1	2	1	2	1	2

Now consider the weighted area enclosed by a path $\pi(\Gamma)$. This is defined as $\sum_c \iota(c)\omega(c)$, where the sum is over all cells in the plane, and $\iota(c)$ is the value of the cell, and $\omega(c)$ is the winding number of $\pi(\Gamma)$ around the cell c. Call this signed value $\rho(\Gamma)$. Notice that if Γ is t_3 or t_4, then $\pi(\Gamma)$ encloses no cells, so $\rho(\Gamma) = 0$. If Γ is t_1, then $\rho(\Gamma) \equiv 1 \pmod 4$, and if Γ is t_2, then $\rho(\Gamma) \equiv -1 \pmod 4$. Hence if τ is a tiling of a region Γ, then $\rho(\Gamma)$ is the sum of the values of $\rho(t)$ for each tile t appearing in τ. Thus $\rho(\Gamma) \equiv a_1(\tau) - a_2(\tau) \pmod 4$. Hence $a_1(\tau) - a_2(\tau) \pmod 4$ is a constant which does not depend on τ. ∎

9.4 Skew tetrominoes and square tetrominoes

In [23], Propp actually considers the tile set consisting of the four skew tetrominoes and the square tetromino. Interestingly, he makes a distinction between two different kinds of squares, based upon the parity of their location. By doing this, he is able to prove a tile invariant which would not have appeared otherwise, namely that the difference between the number of squares of each type is a constant mod ∞. Thus he is able to use this invariant to prove the impossibility of tiling certain regions with the skew tetrominoes alone (by exhibiting a tiling of the region which uses an unequal number of the two types of squares).

He also mentions that there might be a local move property for this set of tiles. (Of course, the distinction between the different types of squares has no bearing on local connectivity.) We will show that no local move property holds for simply connected regions.

Theorem 9.5 *The set of skew and square tetrominoes does not have a local move property for $\mathcal{R}_{sc}$.*

Proof: Consider the region shown in Figure 9-8. This region admits only the two tilings shown, hence the whole region must be one local move. There is an infinite family of regions like this one which must be local moves, hence no local move property holds. ■

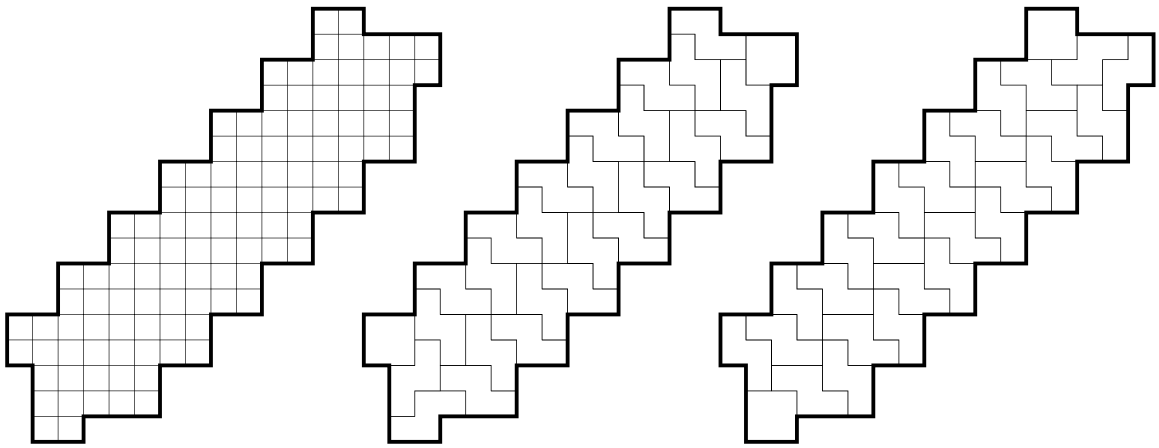


Figure 9-8: A region which can be tiled in only two ways.

Chapter 10

A new approach to ribbon tiles

10.1 Introduction

In this section, we provide a new method for proving the tile invariants for the set of ribbon tiles of order n.

A *ribbon tile of order* n is a polyomino consisting of n squares arranged so that each square is either east or north of the square preceding it. Figure 10-1 shows a ribbon tile of order 11. In general, there will be 2^{n-1} different ribbon tiles of order n. Each one may be indexed by a binary string of length $n-1$ as follows. Begin with the southwesternmost square. Proceed along the tile, writing a 0 each time you move east, and writing a 1 each time you move north. For example, the tile shown in Figure 10-1 would be indexed by the string 0010110001. If t is a ribbon tile, let $\epsilon(t)$ be this binary string. Conversely, if ϵ is a binary string, let t_ϵ denote the corresponding tile. Let ϵ_i denote the ith letter of ϵ, and let $\epsilon_i(t)$ denote the ith letter of $\epsilon(t)$. We will interpret such subscripts mod n. Notice that ϵ_0 is undefined, since the length of ϵ is only $n-1$.

Let $\mathcal{T}_n$ denote the set of all ribbon tiles of order n.

The main result of this section is the following theorem, which was conjectured by Pak in [19], and was first proved in [17].

Theorem 10.1 *Let* Γ *be a simply connected region, let* τ *be a tiling of* Γ *using tiles*

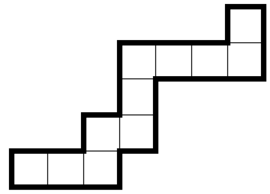


Figure 10-1: The ribbon tile $t_{00101100001}$.

in $\mathcal{T}_n$, and let c be any non-zero integer mod n. Let $a_\epsilon(\tau)$ denote the number of appearances of the tile t_ϵ in the tiling τ. Then

$$\sum_{\epsilon:\epsilon_c=0 \text{ and } \epsilon_{-c}=1} a_\epsilon(\tau) - \sum_{\epsilon:\epsilon_c=1 \text{ and } \epsilon_{-c}=0} a_\epsilon(\tau) \tag{10.1}$$

is a constant (i.e., it depends only on Γ and not on τ).

The approach we take here is similar to the one used in [17]. In section 10.2 we will define boundary words as in [17], observing that the boundary word of each tile corresponds to a closed loop in n dimensions. The new idea we use here is to consider the projection of these loops onto two-dimensional subspaces spanned by pairs of basis vectors. Taking the signed area of these projections will yield the tile invariants.

10.2 Boundary words

Fix an integer n, and consider the set of tiles $\mathcal{T}_n$. Let us adorn the edges of the unit grid as follows. Direct each edge so that it points either south or east. Now for each edge, let (i, j) be the coordinates of its southernmost or westernmost point. Then assign this edge the label k, where $k \equiv i + j \pmod{n}$.

Now for any simply connected region Γ, we may write the boundary word $w(\Gamma)$ of that region by writing the letter z_k whenever we travese an edge labelled k in the direction of that edge, and writing z_k^{-1} whenever we traverse an edge in the opposite direction. For instance, for the region shown in Figure 10-2, the boundary word would be

$$w(\Gamma) = z_2 z_3 z_0 z_1^{-1} z_2^{-1} z_2^{-1} z_2^{-1} z_2^{-1} z_1 z_0 z_3^{-1} z_2.$$

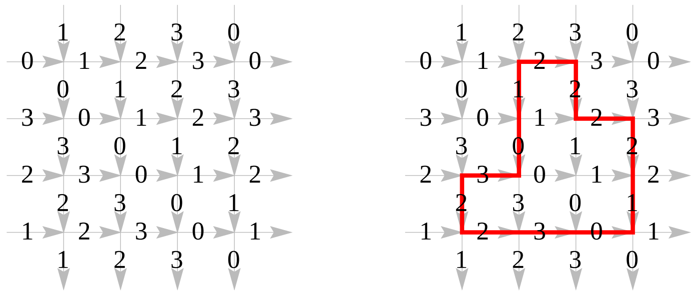


Figure 10-2: An example of boundary words. Here $n = 4$.

Consider the boundary word of a tile $t \in \mathcal{T}_n$. Suppose the southwest corner of the tile lies at a point (x, y), with $x + y \equiv k \pmod{n}$. Let us call this value k the *phase* of t, which we will denote $\phi(t)$. Then the boundary word would be

$$w(t) = z_k z_{k+1}^{\sigma_1} z_{k+2}^{\sigma_2} z_{k+3}^{\sigma_3} \cdots z_{k-1}^{\sigma_{n-1}} z_k^{-1} z_k^{-1} z_{k-1}^{-\sigma_{n-1}} z_{k-2}^{-\sigma_{n-2}} \cdots z_{k+1}^{-\sigma_1} z_k,$$

where $\sigma_c = 1$ if $\epsilon_c(t) = 0$, and $\sigma_c = -1$ if $\epsilon_c(t) = 1$. For example, the order-4 ribbon tile t shown in Figure 10-3 has $\epsilon(t) = 010$, and its boundary word is

$$z_2 z_3 z_0^{-1} z_1 z_2^{-1} z_2^{-1} z_1^{-1} z_0 z_3^{-1} z_2.$$

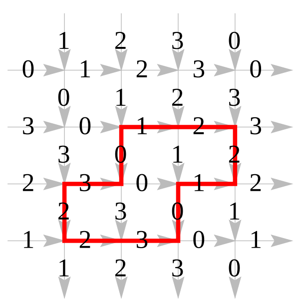


Figure 10-3: Computing the boundary word of an order-4 ribbon tile.

The thing to notice about such a word is that it contains an equal number of occurrences of z_i and z_i^{-1} for every i. We can map such a word onto a closed path in n dimensional space by taking a step in the positive x_i direction wherever a z_i

appears, and taking a step in the negative x_i direction wherever a z_i^{-1} appears.

Lemma 10.2 *Fix an integer n. If Γ is a simply connected region which is tileable by tiles in $\mathcal{T}_n$, then the boundary word will correspond to a closed path in n dimensional space.*

Proof: The boundary word of Γ can be written as the concatenation of conjugates of the boundary words of the individual tiles. The boundary words of each of the tiles form a closed path, so it follows that the boundary word of Γ will be a closed path, too. ■

The converse is not true, however. For $n = 2$, the set $\mathcal{T}_2$ is just the set of dominoes. The region shown in Figure 10-4 has boundary word

$$z_0 z_1^{-1} z_0 z_1 z_0^{-1} z_0^{-1} z_0^{-1} z_0^{-1} z_1 z_0^{-1} z_1^{-1} z_0 z_0 z_0$$

which corresponds to a closed path in 2 dimensions, but it is not tileable by dominoes.

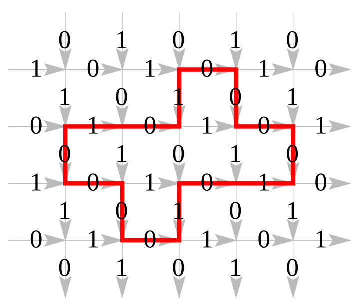


Figure 10-4: A region whose boundary word forms a closed path, but is not tileable. Here $n = 2$.

10.3 Projections

In order to compute the tile invariants, we will take projections of this n-dimensional path onto 2-dimensional subspaces.

For any tileable region Γ, let $\pi(\Gamma)$ denote the closed n-dimensional path that corresponds to its boundary word. Let i and j be distinct integers mod n. Let $\pi_{i,j}(\Gamma)$

denote the projection of $\pi(\Gamma)$ onto the ij-plane. Also, let $w_{i,j}(\Gamma)$ denote the word formed from $w(\Gamma)$ by deleting every symbol which is not z_i, z_j, z_i^{-1}, or z_j^{-1}. We get $\pi_{i,j}(\Gamma)$ from $w_{i,j}(\Gamma)$ by moving right at every occurrence of z_i, left at every occurrence of z_i^{-1}, up at every occurrence of z_j, and down at every occurrence of z_j^{-1}.

The way we will use these 2-dimensional paths is by considering their signed area. For each cell c, the path has a winding number with respect to c, which represents the number of times the path winds around c counterclockwise (minus the number of times it goes around clockwise). The signed area is the sum of the winding numbers of all cells in the 2-dimensional grid. (It is clear that all but finitely many of these winding numbers will be 0.) If the path does not cross itself, then the signed area is precisely the area enclosed (but negated if the path moves clockwise instead of counterclockwise). Let $\alpha_{i,j}(\Gamma)$ denote the signed area of $\pi_{i,j}(\Gamma)$.

Lemma 10.3 *Let t be an order-n ribbon tile located anywhere on the unit grid, and let i and j be distinct integers mod n. Recall that $\phi(t)$ is the phase of t, defined to be the sum of the coordinates of the southwesternmost point of t, taken mod n. Then*

$$\alpha_{i,j}(t) = \begin{cases} 2 & \text{if} \quad \phi(t) = i \quad \text{and} \quad \epsilon_{j-i}(t) = 0 \\ -2 & \text{if} \quad \phi(t) = i \quad \text{and} \quad \epsilon_{j-i}(t) = 1 \\ -2 & \text{if} \quad \phi(t) = j \quad \text{and} \quad \epsilon_{i-j}(t) = 0 \\ 2 & \text{if} \quad \phi(t) = j \quad \text{and} \quad \epsilon_{i-j}(t) = 1 \\ 0 & \text{otherwise} \end{cases}$$

Proof: In section 10.2, we observed that

$$w(t) = z_k z_{k+1}^{\sigma_1} z_{k+2}^{\sigma_2} z_{k+3}^{\sigma_3} \cdots z_{k-1}^{\sigma_{n-1}} z_k^{-1} z_k^{-1} z_{k-1}^{-\sigma_{n-1}} z_{k-2}^{-\sigma_{n-2}} \cdots z_{k+1}^{-\sigma_1} z_k,$$

where $k = \phi(t)$, and where $\sigma_c = 1$ if $\epsilon_c(t) = 0$, and $\sigma_c = -1$ if $\epsilon_c(t) = 1$.

First suppose $\phi(t) = i$. Then

$$w_{i,j}(\Gamma) = z_i z_j^{\sigma_{j-i}} z_i^{-1} z_i^{-1} z_j^{-\sigma_{j-i}} z_i.$$

If $\epsilon_{j-i}(t) = 0$, then $\sigma_{j-i} = 1$, so the path $\pi_{i,j}(t)$ will be a 1×2 rectangle traced counterclockwise, hence $\alpha_{i,j}(t)$ will be 2. Conversely, if $\epsilon_{j-i}(t) = 1$, then $\sigma_{j-i} = -1$, so the path $\pi_{i,j}(t)$ will be a 1×2 rectangle traced clockwise, hence $\alpha_{i,j}(t)$ will be -2.

Now suppose $\phi(t) = j$. Then

$$w_{i,j}(\Gamma) = z_j z_i^{\sigma_{i-j}} z_j^{-1} z_j^{-1} z_i^{-\sigma_{i-j}} z_j.$$

If $\epsilon_{i-j}(t) = 0$, then $\sigma_{i-j} = 1$, so the path $\pi_{i,j}(t)$ will be a 2×1 rectangle traced clockwise, hence $\alpha_{i,j}(t)$ will be -2. Conversely, if $\epsilon_{i-j}(t) = 1$, then $\sigma_{i-j} = -1$, so the path $\pi_{i,j}(t)$ will be a 2×1 rectangle traced counterclockwise, hence $\alpha_{i,j}(t)$ will be 2.

If $\phi(t) \neq i$ and $\phi(t) \neq j$, then

$$w_{i,j}(\Gamma) = z_i^{\sigma_{i-k}} z_j^{\sigma_{j-k}} z_j^{-\sigma_{j-k}} z_i^{-\sigma_{i-k}} \quad \text{or} \quad w_{i,j}(\Gamma) = z_j^{\sigma_{j-k}} z_i^{\sigma_{i-k}} z_i^{-\sigma_{i-k}} z_j^{-\sigma_{j-k}}$$

depending on the relative cyclic order of i, j, and k, where $k = \phi(t)$. In either case, the path $\pi_{i,j}(t)$ consists of four steps which form an L shape and then retrace their path. Thus the signed area $\alpha_{i,j}(t)$ will be 0. ■

Unfortunately, the function $\alpha_{i,j}(t)$ depends on the phase of the tile t. In order to prove tile invariants, we will need a function which does not depend on $\phi(t)$. The following definition will do the trick. Let c be any non-zero integer mod n. Then define

$$\xi_c(\Gamma) = \sum_{a=0}^{n-1} \alpha_{a,a+c}(\Gamma).$$

Lemma 10.4 *Let t be an order-n ribbon tile located anywhere on the unit grid, and let c be a non-zero integer mod n. Then*

$$\xi_c(t) = \begin{cases} 4 & \text{if } \epsilon_c(t) = 0 \text{ and } \epsilon_{-c}(t) = 1 \\ -4 & \text{if } \epsilon_c(t) = 1 \text{ and } \epsilon_{-c}(t) = 0 \\ 0 & \text{otherwise} \end{cases}$$

Proof: Let $k = \phi(t)$. The only terms in the summation which will be non-zero are those for which $a = k$ and $a = k - c$. Thus

$$\xi_c(t) = \alpha_{k,k+c}(t) + \alpha_{k-c,k}(t).$$

From Lemma 10.3, we have that $\alpha_{k,k+c}(t) = 2$ if $\epsilon_c(t) = 0$, and $\alpha_{k,k+c}(t) = -2$ if $\epsilon_c(t) = 1$. Similarly, $\alpha_{k-c,k}(t) = -2$ if $\epsilon_{-c}(t) = 0$, and $\alpha_{k-c,k}(t) = 2$ if $\epsilon_{-c}(t) = 1$. The result now follows easily. ∎

We can now prove Theorem 10.1. Let Γ be a simply connected region, and let τ be a tiling of Γ. Observe that $\xi_c(\Gamma)$ is equal to $\sum_{t\in\tau} \xi_c(t)$, where we sum over all tiles in τ. By Lemma 10.4, this is precisely equal to the expression (10.1), times 4. Hence (10.1) is equal to $\frac{1}{4}\xi_c(\Gamma)$, which does not depend on τ. This proves Theorem 10.1.

Chapter 11

Augmenting untileable regions

11.1 Introduction

Let us consider once again the topic of domino tilings. Let Γ be a simply connected region. Color the cells of the unit grid blue and gold in checkerboard fashion. Let us say that the region Γ is *balanced* if it contains an equal number of blue and gold cells. As we have noted before, being balanced is a necessary condition for tileability, but it is not sufficient, as illustrated in Figure 11-1.

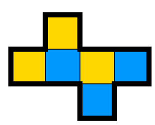


Figure 11-1: A balanced region which is not tileable.

However, in this case, if we enlarge the region by adding a domino shape to the outside of the region, we can get a region which is tileable, as shown in Figure 11-2.

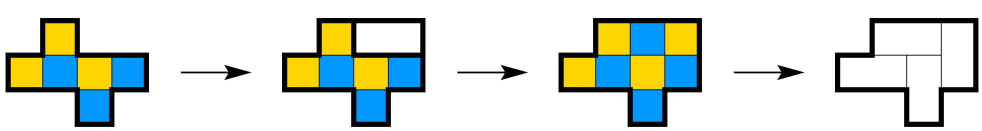


Figure 11-2: Adding a domino to the outside of Γ yields a larger region which is tileable.

In this case, we were able to add dominoes to the outside of Γ to make Γ tileable.

In the next section, we will prove that this is always possible for a particular class of regions. Then in section 11.3, we will see an example where this is not true.

11.2 Row-convex regions

Let Γ be a simply-connected region. We say that Γ is *row-convex* if every horizontal line intersects Γ in an interval. In other words, if c_1 and c_2 are cells of Γ which lie in the same row, then all cells between c_1 and c_2 must also lie in Γ.

Theorem 11.1 *Let Γ be a balanced, row-convex region of the plane. Then there exists a row-convex region R containing Γ such that $R - \Gamma$ is tileable by dominoes and R is tileable by dominoes.*

Proof: Our proof consists of two steps. First we will show that any row-convex region Γ can be extended to a so-called *trapezoidal* region R by adding dominoes. Then we will show that all balanced trapezoidal regions are domino-tileable.

We say a region R is *trapezoidal* if every cell in R which is not in the top row of R has a cell immediately above it. Figure 11-3 shows a trapezoidal region.

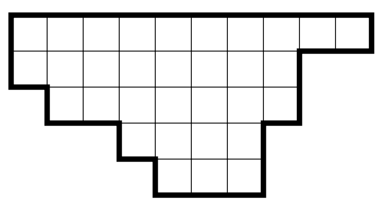


Figure 11-3: A trapezoidal region.

Take a region Γ which is not trapezoidal. Say that a row r of Γ is *satisfactory* if every cell of Γ in the row below r has a cell directly above it. Take the lowest row which is not satisfactory. Add horizontal dominoes to either end of this row until it is satisfactory. Do this for each row until every row is satisfactory. The resulting region is then trapezoidal. See Figure 11-4.

Clearly if the original region Γ is balanced, the augmented region R will be balanced as well. Now we must show that all balanced trapezoidal regions are tileable by dominoes. We prove this by induction on the size of the region.

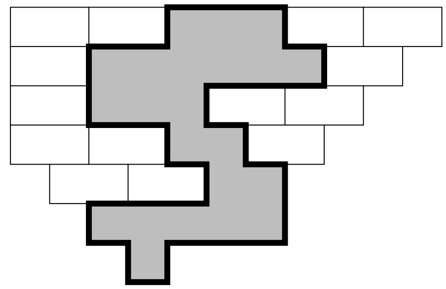


Figure 11-4: Adding dominoes to create a trapezoidal region.

A region of size zero is trivially tileable. Suppose R is a balanced trapezoidal region with area $2k$. It will suffice to show the existence of a 2×1 rectangle D contained in R such that $R - D$ is trapezoidal.

Consider the southeast boundary af R, running from the lower-left corner of the bottom row to the top right corner of the top row. The boundary consists of horizontal and vertical segments. If it contains a horizontal segment of length at least 2, we can remove the domino as shown in the first panel of Figure 11-5. If it contains a vertical segment of length at least 2, we can remove the domino as shown in the second panel of Figure 11-5. So we may assume this boundary consists of alternating horizontal and vertical segments of length 1. The same holds for the southwest boundary. So R must be a triangle like the one shown in Figure 11-6. But this is impossible, since such a figure is not balanced.

Thus all balanced trapezoidal regions are tileable, proving the theorem. ∎

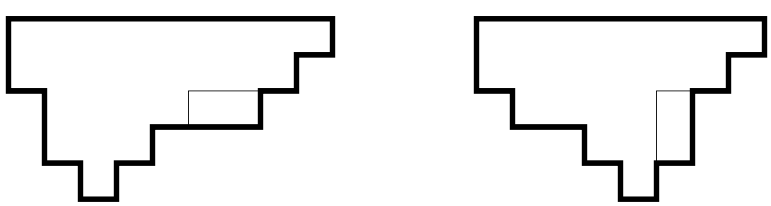


Figure 11-5: Removing a domino from a trapezoidal region.

11.3 Regions where this fails

It turns out that this theorem fails for regions which are not row-convex.

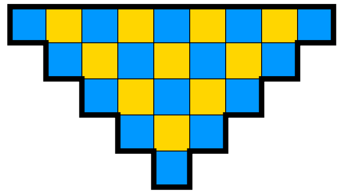


Figure 11-6: A triangle.

Proposition 11.2 *For the balanced region Γ shown in Figure 11-7, there is no region R which contains it such that both R and $R - \Gamma$ are tileable by dominoes.*

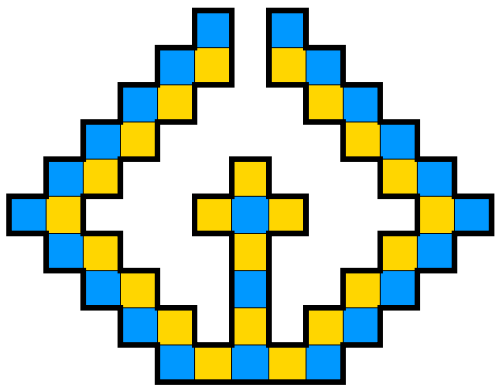


Figure 11-7: A balanced region which cannot be augmented to a tileable region.

Proof: Let us consider the problem in its dual form. Think of each cell of the square grid as a vertex. Now Γ is an induced subgraph of the infinite grid. The question is whether there exists an induced subgraph R which contains Γ such that both R and $R - \Gamma$ have perfect matchings. Suppose such an R exists. Then there exist perfect matchings of both R and $R - \Gamma$. Direct edges of the first matching so they go from blue vertices to gold vertices, and direct edges of the second matching so that they go from gold vertices to blue vertices. Now consider the union of these two matchings.

Every vertex in $R - \Gamma$ is incident to one edge from each matching, so such a vertex has indegree 1 and outdegree 1. Every blue vertex in Γ has outdegree 1 and indegree 0, while every gold vertex in Γ has indegree 1 and outdegree 0. Thus, this union of matchings must consist of disjoint paths from blue vertices of Γ to gold vertices of Γ, plus possibly some cycles lying outside of Γ. See Figure 11-8.

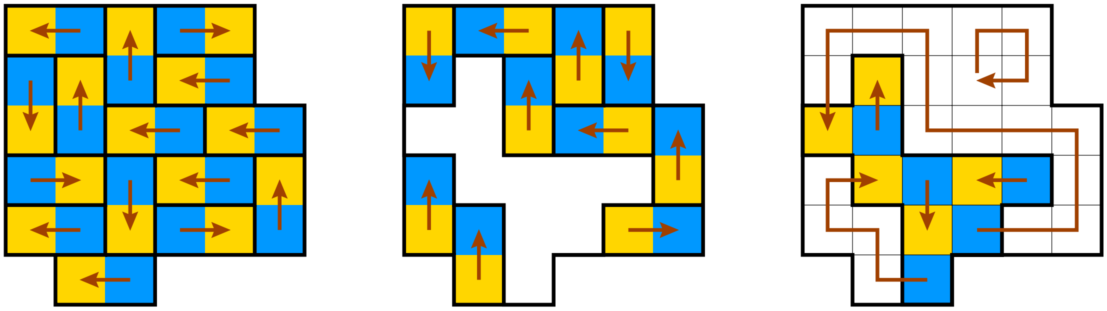


Figure 11-8: The union of the brown matchings forms disjoint paths from blue vertices of Γ to gold vertices of Γ.

For a region Γ, define the graph G_Γ as follows. Include a vertex for every cell of the square grid (both those squares in Γ and not in Γ), connecting those cells which are adjacent in the grid. (So far, G_Γ is just an infinite square grid.) Now add two distinguished vertices, call them V_{blue} and V_{gold}. Connect V_{blue} to every blue cell in Γ, and connect V_{gold} to every gold cell in Γ.

Suppose there exists a region R containing Γ such that both R and $R - \Gamma$ are domino-tileable. Then there must exist k vertex-disjoint paths from V_{blue} to V_{gold} in G_Γ, where k is the number of blue vertices (or gold vertices) in Γ.

At this point, let us recall Menger's Theorem, in one of its many forms.

Theorem 11.3 *(Menger's Theorem.) Let G be an undirected graph, let s and t be non-adjacent vertices of G. Then the maximum number of internally-vertex-disjoint paths from s to t equals the minimum number of vertices from $V(G) - \{s, t\}$ whose deletion separates s and t.*

Now let us apply this to our specific example. If there exists a region R containing Γ such that both R and $R - \Gamma$ are domino-tileable, then there must exist 23 vertex-disjoint paths from V_{blue} to V_{gold} in G_Γ. So it will suffice to find 22 vertices of G_Γ whose deletion disconnects V_{blue} and V_{gold}. Figure 11-9 shows such a set. This proves Proposition 11.2. ∎

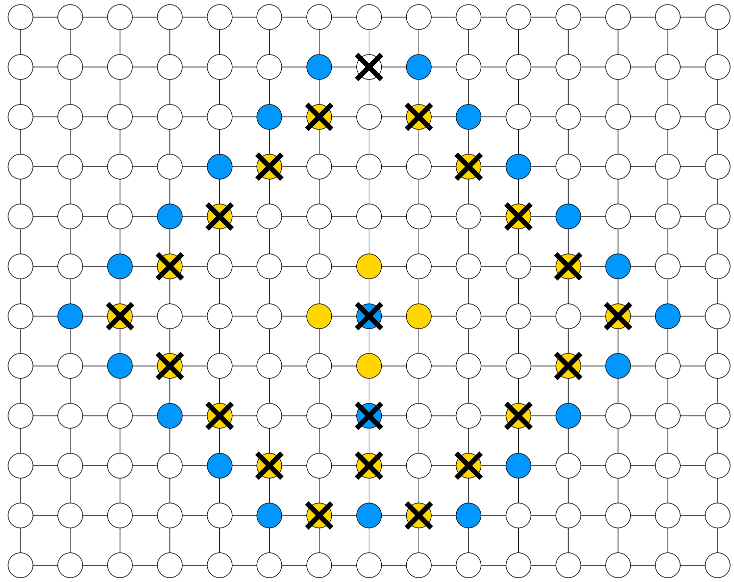


Figure 11-9: Deleting the 22 vertices marked with X's eliminates all paths from blue vertices to gold vertices, hence it separates the distinguished vertices V_{blue} and V_{gold}.

11.4 Three or more dimensions

Theorem 11.4 *Theorem 11.1 holds for balanced row-convex regions in n dimensions.*

Proof: Label each cell of the region with coordinates $(x_1, x_2, \ldots, x_n)$. By row-convex, we mean that if $a < b < c$, and $(a, x_2, \ldots, x_n)$ and $(c, x_2, \ldots, x_n)$ are cells of Γ, then $(b, x_2, \ldots, x_n)$ is a cell of Γ as well.

Let $y_1, \ldots, y_n, z_1, \ldots, z_n$ be values such that if $(x_1, \ldots, x_n) \in \Gamma$, then $y_i \leq x_i \leq z_i$ for all i. (In other words, the box bounded by these values contains the region Γ.) Consider the $(n-1)$-dimensional box B given by the inequalities $y_i \leq x_i \leq z_i$ for $2 \leq i \leq n$. Let $\alpha : \mathbb{Z} \to \mathbb{Z}^{n-1}$ be a Hamiltonian path through the cells of B. (Specifically, for $1 \leq j \leq |B|$, each $(n-1)$-tuple $\alpha(j)$ is a different cell of B, and $\alpha(j)$ and $\alpha(j+1)$ are neighboring cells.)

Consider the map $\phi : \mathbb{Z} \times B \to \mathbb{Z}^2$ defined as follows:

$$\phi(x_1, x_2, \ldots, x_n) = (x_1, \alpha^{-1}(x_2, \ldots, x_n)).$$

Observe the following facts about ϕ. The map ϕ is a bijection between $\mathbb{Z} \times B$ and $\mathbb{Z} \times [1, |B|]$. The image of a balanced region is balanced. The image of a row-convex region is row-convex (though not necessarily connected). If c_1 and c_2 are neighboring cells in $\mathbb{Z} \times [1, |B|]$, then $\phi^{-1}(c_1)$ and $\phi^{-1}(c_2)$ are neighboring cells of $\mathbb{Z} \times B$. In particular this means that if $R \subset \mathbb{Z} \times [1, |B|]$ is tileable by dominoes, then $\phi^{-1}(R)$ is tileable by dominoes.

Recall that Γ is our balanced row-convex n-dimensional region. Consider $\phi(\Gamma)$. By the above observations, $\phi(\Gamma)$ is a balanced, row-convex 2-dimensional region. So by Theorem 11.1, there exists a region R containing $\phi(\Gamma)$ such that both R and $R - \phi(\Gamma)$ are domino-tileable. (Notice that the proof of Theorem 11.1 does not require that the original region be connected. Also notice that R will be no taller than $\phi(\Gamma)$, so if $\phi(\Gamma) \subset \mathbb{Z} \times [1, |B|]$, then $R \subset \mathbb{Z} \times [1, |B|]$ as well.) So $\phi^{-1}(R)$ is an n-dimensional region containing Γ. Then $\phi^{-1}(R)$ is domino-tileable since R is, and $\phi^{-1}(R) - \Gamma$ is domino-tileable since $R - \phi(\Gamma)$ is.

This proves the theorem. ■

Chapter 12

A counterexample in three dimensions

12.1 Introduction

One of the techniques we used to attack various tiling problems in two dimensions was the notion of boundary words. One of the key observations that makes the technique work is that the boundary word of a region can be written as the product of conjugates of the boundary words of the tiles which tile the region. In order to prove this fact, one needs to use the fact that in any tiling τ, there exists a tile which, when deleted from Γ, yields a region which is still simply connected. This is fairly straightforward to see in two dimensions.

In [18], Pak considers whether it could be possible to use a similar technique in three dimensions. He observes that the analogous fact in three dimensions does not hold, and he cites as an example a set of six pieces which can be fitted together to form a shape which is topologically equivalent to a ball, but such that removal of any one piece leaves a hole in the shape (Figure 12-1). The problem with this example, he says, is that it is possible to take a subset of the pieces (consisting of three pieces), such that the subset and its complement each form shapes topologically equivalent to a ball.

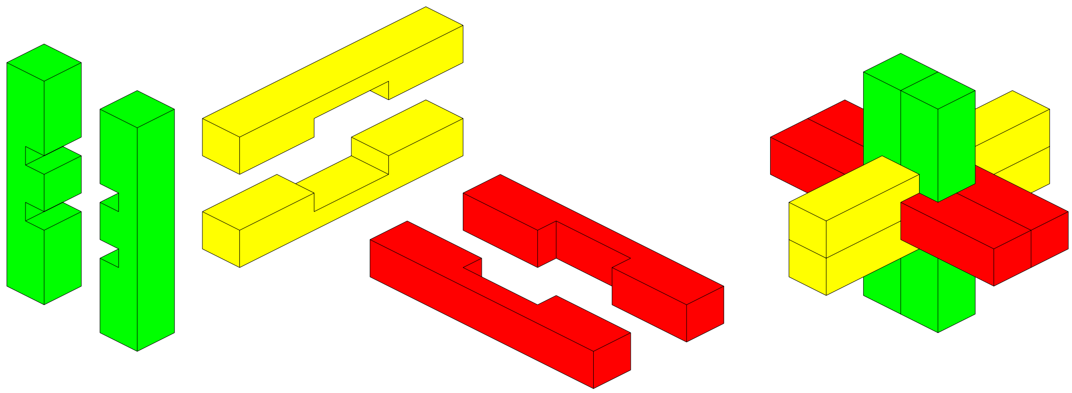


Figure 12-1: Six pieces which can be fitted together to form a shape topologically equivalent to a ball.

12.2 The counterexample

In the following picture (Figure 12-2), we have three solid shapes which can be fitted together to form a rectangular box. (In other words, they tile the box in three dimensions.) The original shapes and the box are all topologically equivalent to a ball. However, when we delete any one of the three shapes from the tiling, the resulting shape is not topologically equivalent to a ball. Thus this tiling cannot be divided into two subsets each of which is equivalent to a ball.

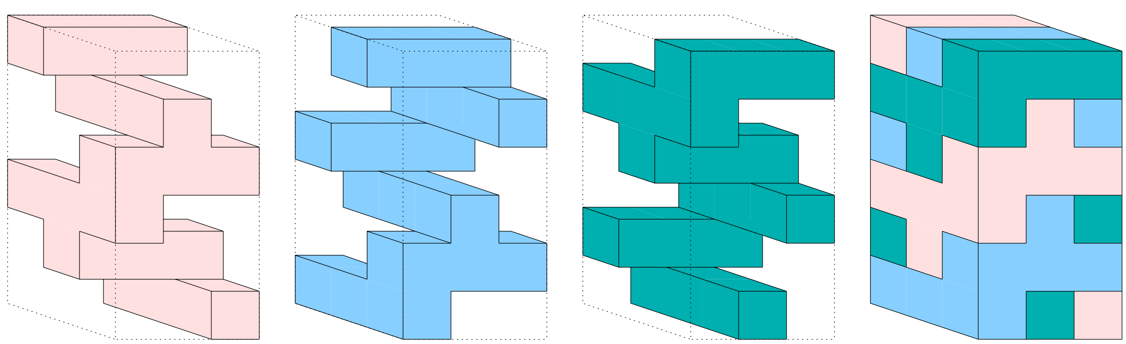


Figure 12-2: Three shapes which fit together to tile a box.

Bibliography

[1] D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs.* monograph in preparation.

[2] O. Bodini and E. Rémila. Tiling a polygon with two kinds of rectangles. 2002.

[3] William M. Boyce. Generation of a class of permutations associated with commuting functions. *Math. Algorithms 2 (1967), 19–26; addendum, ibid.*, 3:25–26, 1967.

[4] David M. Bressoud. *Proofs and confirmations.* MAA Spectrum. Mathematical Association of America, Washington, DC, 1999. The story of the alternating sign matrix conjecture.

[5] Hal Canary. Aztec Diamonds and Baxter Permutations, 2004.

[6] F. R. K. Chung, R. L. Graham, V. E. Hoggatt, Jr., and M. Kleiman. The number of Baxter permutations. *J. Combin. Theory Ser. A*, 24(3):382–394, 1978.

[7] J. H. Conway and J. C. Lagarias. Tiling with polyominoes and combinatorial group theory. *J. Combin. Theory Ser. A*, 53(2):183–208, 1990.

[8] Robert Cori, Serge Dulucq, and Gérard Viennot. Shuffle of parenthesis systems and Baxter permutations. *J. Combin. Theory Ser. A*, 43(1):1–22, 1986.

[9] Noam Elkies, Greg Kuperberg, Michael Larsen, and James Propp. Alternating-sign matrices and domino tilings. I. *J. Algebraic Combin.*, 1(2):111–132, 1992.

[10] Noam Elkies, Greg Kuperberg, Michael Larsen, and James Propp. Alternating-sign matrices and domino tilings. II. *J. Algebraic Combin.*, 1(3):219–234, 1992.

[11] K. Eloranta. Diamond ice. *J. Statist. Phys.*, 96(5-6):1091–1109, 1999.

[12] P. Erdős and R. L. Graham. On a linear diophantine problem of Frobenius. *Acta Arith.*, 21:399–408, 1972.

[13] S. W. Golomb. *Polyominoes*. Georges Allen and Unwin Ltd, London, 1966.

[14] William Jockusch, James Propp, and Peter Shor. Random Domino Tilings and the Arctic Circle Theorem, 1995.

[15] C. Kenyon and R. Kenyon. Tiling a polygon with rectangles. *Proc. 33-rd FOCS*, pages 610–619, 1992.

[16] K. Makarychev and Y. Makarychev. Proof of Pak's conjecture on tilings by T-tetrominoes (in Russian).

[17] Cristopher Moore and Igor Pak. Ribbon tile invariants from the signed area. *J. Combin. Theory Ser. A*, 98(1):1–16, 2002.

[18] I. Pak. Tile invariants: New horizons. *Theoret. Comput. Sci.*, 303:303–331, 2003.

[19] Igor Pak. Ribbon tile invariants. *Trans. Amer. Math. Soc.*, 352(12):5525–5561 (electronic), 2000.

[20] J. Propp. Generating random elements of finite distributive lattices. *Electron. J. Combin.*, 4(2):Research Paper 15, approx. 12 pp. (electronic), 1997. The Wilf Festschrift (Philadelphia, PA, 1996).

[21] J. Propp and D. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

[22] J. Propp and D. Wilson. Coupling from the past: a user's guide. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 181–192. Amer. Math. Soc., Providence, RI, 1998.

[23] James Propp. A pedestrian approach to a method of Conway, or, A tale of two cities. *Math. Mag.*, 70(5):327–340, 1997.

[24] James Propp. The many faces of alternating-sign matrices. In *Discrete models: combinatorics, computation, and geometry (Paris, 2001)*, Discrete Math. Theor. Comput. Sci. Proc., AA, pages 043–058 (electronic). Maison Inform. Math. Discrèt. (MIMD), Paris, 2001.

[25] Alistair Sinclair. *Algorithms for random generation and counting.* Progress in Theoretical Computer Science. Birkhäuser Boston Inc., Boston, MA, 1993.

[26] Richard P. Stanley. *Enumerative combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1997. With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.

[27] W. Thurston. Conway's tiling groups. *Amer. Math. Monthly*, 97(8):757–773, 1990.

[28] D. W. Walkup. Covering a rectangle with T-tetrominoes. *Amer. Math. Monthly*, 72:986–988, 1965.

[29] D. J. A. Welsh. *Complexity: knots, colourings and counting*, volume 186 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, Cambridge, 1993.

[30] D. West. An elementary proof of two triangle-tiling theorems of Conway and Lagarias. 1991.