HAL open science

Pavages Aléatoires

Alexandra Ugolnikova

▶ To cite this version:

Alexandra Ugolnikova. Pavages Aléatoires. Modélisation et simulation. Université Sorbonne Paris Cité, 2016. Français. NNT : 2016USPCD034 . tel-01889871

HAL Id: tel-01889871

https://theses.hal.science/tel-01889871v1

Submitted on 8 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS NORD (PARIS 13)
Laboratoire d'Informatique de Paris Nord

Thèse

pour obtenir le titre de

Docteur de l'Université Paris Nord (Paris 13)

Spécialité Informatique

PAVAGES ALÉATOIRES

présentée et soutenue publiquement par

Alexandra UGOLNIKOVA

le 2ème décembre 2016

devant le jury comprosé de:

Directeur de thèse : **Thomas FERNIQUE**

Rapporteurs: **Béatrice DE TILIÈRE**
Éric RÉMILA

Examinateurs: **Frédérique BASSINO**
Olivier BODINI
Ana BUŠIĆ
Pavel KALOUGUINE

Abstract

In this thesis we study two types of tilings: tilings by a pair of squares and tilings on the tri-hexagonal (Kagome) lattice. We consider different combinatorial and probabilistic problems. First, we study the case of 1×1 and 2×2 squares on infinite stripes of height k and get combinatorial results on proportions of 1×1 squares for $k \leq 10$ in plain and cylindrical cases. We generalize the problem for bigger squares. We consider questions about sampling and approximate counting. In order to get a random sample, we define Markov chains for square and Kagome tilings. We show ergodicity and find polynomial bounds on the mixing time for $n \times \log n$ regions in the case of tilings by 1×1 and $s \times s$ squares and for lozenge regions in the case of restrained Kagome tilings. We also consider weighted Markov chains where weights λ are put on the tiles. We show rapid mixing with conditions on λ for square tilings by 1×1 and $s \times s$ squares and for Kagome tilings. We provide simulations that suggest different conjectures, one of which existence of frozen regions in random tilings by squares and on the Kagome lattice of regions with non flat boundaries.

Résumé

Dans cette thèse nous étudions deux types de pavages: des pavages par une paire de carrés et des pavages sur le réseau tri-hexagonal (Kagome). Nous considérons différents problèmes combinatoires et probabilistes. Nous commencons par le cas des carrés 1×1 et 2×2 sur des bandes infinies de hauteur k et obtenons des resultats sur la proportion moyenne des carrés 1×1 pour les cas planaire et cylindrique pour $k \leq 10$. Nous considerons également des questions d'échantillonnage et comptage approximatif. Pour obtenir un échantillon aléatoire nous définissons des chaînes de Markov pour les pavages par des carrés et sur le réseau Kagome. Nous montrons des bornes polynomiales pour le temps de mélange pour les pavages par des carrés 1×1 et $s \times s$ des régions $n \times \log n$ et les pavages Kagome des régions en forme de losange. Nous considérons aussi des chaines de Markov avec des poids λ sur les tuiles. Nous montrons le mélange rapide avec des conditions spécifiques sur λ pour les pavages par des carrés 1×1 et $s \times s$ et pavages Kagome. Nous présentons des simulations qui suggèrent plusieurs conjectures, notamment l'existence des région gelées pour les pavages aléatoires par des carrés et sur le réseau Kagome des régions avec des bords non plats.

Acknowledgements

First of all, I would like to thank my referees, Béatrice de Tilière and Éric Rèmila for agreeing for reading this manuscript and writing reviews. I would like to thank Frédérique Bassino, Olivier Bodini, Ana Busic and Pavel Kalouguine for agreeing to be part of the jury committee.

There are so many people who have provided support throughout these years. I would like to thank all of you for your help, whether it was academic help, administrative work or moral support during the time of need. I would like to thank my supervisor, Thomas Fernique, for introducing me to the world of tilings, for spending hours explaining various things, for teaching me how to make inskape pictures, for making me teach programming classes, for reading drafts of this thesis and of course for support that he provided during these three years. I would like to thank again Olivier Bodini for being my official supervisor and for suggesting a great square problem 2,5 years ago that I have been struggling to solve ever since.

I would like to thank Dana Randall for being an inspiration in the field of Markov chain mixing, Pavel Kalouguine for patient explanations about counting problems in statistical physics, Johan Nilsson and Nicolas Rolin for our long discussions about big and small squares. I would also like to thank Iulia Beloshapka who has always been an inspiration to me and without whom I would have never started this doctoral program and Quentin Marcou for a cython crash course and for useful remarks about the final version of the text.

I would like to thank all CALIN team of Laboratoire d'Informatique de Paris Nord for providing a great working environment, my friends and family for an incredible amount of moral support. I trust no one will be offended for my failure to name them.

Contents

Abstract ii

Acknowledgements v

Introduction 1

1 Dimer tilings 7
- 1.1 Introduction . . . 7
- 1.2 Counting dimers . . . 9
  - 1.2.1 Adjacency matrices . . . 9
  - 1.2.2 Plane partitions . . . 10
  - 1.2.3 Aztec diamond . . . 10
  - 1.2.4 Inverse Kasteleyn matrix . . . 11
  - 1.2.5 Entropy . . . 12
- 1.3 Random generation . . . 13
  - 1.3.1 Height function . . . 14
  - 1.3.2 Classical Markov chain MC_{simple} . . . 16
  - 1.3.3 Modified Markov chain MC_{tower} . . . 18
  - 1.3.4 Back to single flips . . . 20
  - 1.3.5 Domino shuffling . . . 20
  - 1.3.6 Limit shape . . . 21

2 Square tilings 23
- 2.1 Introduction . . . 23
- 2.2 Combinatorics of $(1,2)-$square tilings . . . 24
  - 2.2.1 Bivariate generating function . . . 25
  - 2.2.2 Counting bounds . . . 26
  - 2.2.3 Combinatorial results . . . 29
  - 2.2.4 Automaton representation . . . 30
    - 2.2.4.1 Essential, non-essential and additional states . . . 31
    - 2.2.4.2 Simplified automata . . . 34
  - 2.2.5 Matrix representation . . . 36
  - 2.2.6 Cylindrical case . . . 37
  - 2.2.7 Open problems . . . 39
- 2.3 Combinatorics of $(1,s)-$square tilings . . . 39
  - 2.3.1 Comparison . . . 40
  - 2.3.2 1-dimensional case . . . 43
- 2.4 Topological entropy and approximate counting . . . 45

2.5 Random generation . . . 46
2.5.1 Height function . . . 47
2.5.1.1 Quotient group H . . . 47
2.5.1.2 Weighted Cayley graph . . . 48
2.5.1.3 Height function for $(m, s)-$square tilings . . . 50
2.5.1.4 Flips . . . 51
2.5.2 Examples . . . 52
2.5.3 Markov chains . . . 54
2.5.4 Mixing time for $(1, s)-$square tilings . . . 55
2.5.5 Weighted Glauber dynamics . . . 58
2.5.5.1 $(1, 2)-$square tilings with weights . . . 59
2.5.5.2 $(1, s)-$square tilings with weights . . . 62
2.5.6 Simulations . . . 63
2.5.7 Limit shape . . . 65

3 Kagome tilings **67**
3.1 Introduction . . . 67
3.2 Approximate counting . . . 68
3.2.1 Kagome tilings as perfect matching . . . 68
3.2.2 Perfect matchings in bipartite graphs . . . 70
3.2.3 Back to Kagome . . . 71
3.3 Random generation . . . 72
3.3.1 Height function . . . 72
3.3.2 Flips . . . 73
3.3.3 Markov chain . . . 75
3.3.4 Coupling of general Markov chain . . . 75
3.3.5 Weighted Glauber dynamics . . . 76
3.3.6 Restrained Kagome tilings . . . 78
3.3.7 Limit shape . . . 83

Conclusion **83**

Appendices **89**

A Markov chains and mixing time **91**
A.1 Markov chain . . . 91
A.2 Mixing and Coupling times . . . 93

B Random generation **97**
B.1 Approximate counting . . . 97
B.2 Grand coupling or sanwiching . . . 98
B.3 Coupling from the Past . . . 99

Bibliography **101**

Introduction

A **tiling** (or **tessellation**) is a covering of a region by given geometric figures (called **tiles**) with no holes or overlapping. Mathematical foundation and technical terms used when discussing tilings can be found in [35]. Here we consider tilings on regular lattices where corners of tiles coincide with the vertices of the lattice (See Figure 1). A **random** tiling is a tiling that is chosen uniformly at random from a set of tilings of a given region.

Tilings have attracted interest and attention not only from a scientific point of view but esthetic as well. From the ancient times when the question of tileability was first asked to the late 20th century when tilings started being used to represent the structure of quasicrystals. With the invention of the Turing machine, tilings were found to be an excellent way to represent a Turing machine using Wang tiles. And so the question about running time of the Turing machine could be seen as the tileability problem using Wang tiles.

Tilings have been a helpful tool in crystallography. They are used to model crystals and answer questions about the growth of crystals from the theoretical point of view. Since the discovery of quasicrystals in the 1980s [1] researchers have been trying to understand the way quasicrystals form and grow, what happens with their molecular structure when they cool down or heat up. One of the tiling model associated with quasicrystals are tilings by rhombis that have 5-fold rotational symmetry.

[1] Quasicrystals are crystals that have symmetries that are forbidden by usual crystallographic rules (such as 5-fold or 8-fold for example). See more in [28, 42, 88]

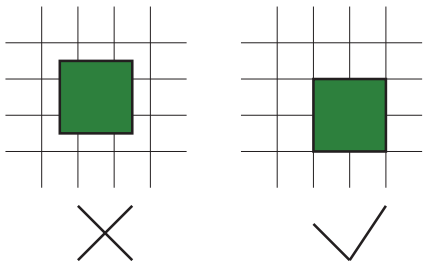

FIGURE 1: Not allowed (left) and allowed (right) placements of a square tile on the $\mathbb{Z}^2$ lattice.

Since tilings can be seen as a model for quasicrystals, one of the questions is to find a theory for the growth process.Tilings with minimal energy are characterized by matching rules that allow some local patterns in a tiling. Then a question that one can ask is: when one start from a tiling with a lot of mismatches, and at each step corrects it by performing local transformation, how long before one arrives to a tiling with minimal energy? This process is often referred to as the relaxation process and was studied in the one-dimensional case for two-letter words and in the two-dimensional case for rhombus tilings by [7, 8, 29].

There are different questions that can be asked in the field of tilings. In statistical mechanics the research is focused on getting some counting results which are usually based on the assumption about the physical system that is under study. There are few known counting results for systems such as dimers on $\mathbb{Z}^2$, triangular lattices, hard hexagons. Various parameters can be added to the tiling system, such as the energy and temperature, to extract some properties about the system such as phase transitions or long range order. Some well-studied models are the Ising model and the Hardcore model.

From a mathematician's point of view, it is interesting to induce some structure on tilings and get rigorous proofs about the number of ways in which a set of tiles can tile a region, properties that a typical tiling has. One of the fundamental works on the structure of tilings is Conway's tiling groups by Thurston [94] which allows to use finite groups as a tool to induce order on the set of tilings. It allows to represent each tile as a relator in a group and to get a necessary condition on the tileability of a region by checking whether the group element representing the boundary of the region is the trivial element of the group.

A computer scientist poses questions related to the complexity of designed algorithms that could check tileability of a region by a given set of tiles or would produce random tilings and approximate schemas for counting problems. The questions of tileability generally belongs to the class of NP-problems. For example, tileability of a finite region by a pair of rectangles is NP-complete [83], while it can be done in quadratic time for a simply connected region [82].

Let us point out that the only tilings that are well studied from all points of view are domino and lozenge tilings. They are also known as dimers on square and triangular lattices or perfect matchings on the square and honeycomb lattices. There are exact counting formulas for the number of domino and lozenge tilings for some specific regions. In general counting problems are hard. For example, the counting problem for perfect matching was proven to be $\#P$ [97]. So it is considered that we cannot efficiently count

the number of perfect matchings exactly. But there is a possibility to get an approximation on the counting problem in polynomial time within a constant [approximation] factor.

Sometimes an approximate counting problem can be reduced to a problem of sampling from the given set of configurations. It is generally done by defining a Markov chain on the state space that includes all configurations from which we wish to sample and that has the desired stationary distribution. Then if we simulate the chain long enough, we can can get a sample from a distribution arbitrary close to the stationary distribution. The stationary distribution is often taken to be uniform, in which case the sample is almost uniform.

The main motivation behind this work is to consider tiling models that are a bit more difficult than dimers. We study two tiling families: square tilings which are tilings by two type of squares of sizes m and s and Kagome tilings – tilings on the Kagome lattice where a tile is an hexagon with two adjacent triangles (see Figure 2). The Kagome lattice[2] itself appears in the atomic crystal structure of some minerals (namely jarosites and herbertsmithite). Tilings on the Kagome lattice were studied by Bodini in [6]. Square tilings were studied as a case of rectangular tilings By R. Kenyon, C. Kenyon and Rémila in [54, 55, 82].

The choice of these particular tiling families is due to the fact that there is an integer-valued function that can be defined on vertices (and edges) of a tiling which allows to define an order on the set of tilings (height function). And more importantly, the tiling space of a simply connected region is connected by local transformations. There might be not one but two or more height functions: this is the case for quadri-tilings[3] which were studied in [96].

Other models for which the configurational space is proven to be connected via local transformations are: T-tetramino tilings by Korn-Pak in [60] and ribbon tilings which appear in [75, 87]. Let us point out that we consider models with a small set of tiles. When the set of tiles is big, different models of dissections come to mind: equitable

[2] It is one of 11 uniform tilings of the Euclidean plane by regular polygons

[3] A quadri-tiling is an edge-to-edge tiling by quadri-tiles. Consider a set of right triangles whose hypotenuses have length 2; moreover for each triangle the endpoints of the hypotenuse are colored in white while the third vertex is colored in black. A quadri-tile is obtained by gluing together two right triangles along the hypotenuse or along an edge if possible respecting the colors.

rectangular dissections [4], dyadic tilings [5] and triangulations [6]. These models were proved to have flip-accessibility and were studied in the sense of random generation. Bounds on the mixing time of the considered weighted Markov chains were obtained for triangulations [13] and for dyadic tilings and rectangular dissections in [11].

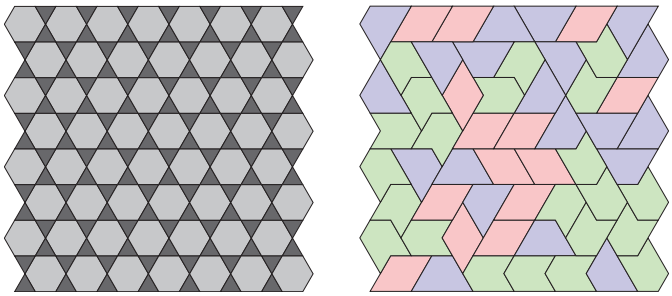

FIGURE 2: Part of the Kagome lattice (left) and a tiling on it (right).

We start with Chapter 1 where we present the main results about dimer tilings which have been a subject of extensive research (some references are: [14, 16, 19, 26, 27, 52, 56, 57, 66, 78]). We mention formulas for counting dimers. We recall how a height function is defined so that dimer tilings can be seen as discrete surfaces in $\mathbb{R}^3$ and and how it is used to calculate the entropy of a random tiling using a variational principle. Then we focus on random generation. We introduce a natural Markov chain and show why the path coupling method does not apply directly. Then we summarize different ways to bound the mixing time of the Markov chain (see [14], [66], [99]). At the end we mention the Artic circle phenomenon (see [16, 26]) which occurs for random domino tilings of the Aztec diamond and for lozenge tilings for hexagonal regions.

Chapter 2 is dedicated to square tilings. First, we consider tilings by 1×1 (small) and 2×2 (big) squares of an infinite strip of height k. We have published these results in [85]. We use combinatorial tools to find average proportion of space occupied by small squares in plain and cylindrical cases for k up tol 10. We conjecture the sequence of average proportions to be converging when $k \to \infty$. Then we generalize the construction for tilings by 1×1 and $s \times s$ squares and present the computational comparison of average proportions of small squares for infinite stripes of height up to 10 and square regions of size up to 30. We show that in the $1-$dimensional case the average proportion of small tiles grows with s. We conjecture that it is the same in the $2-$dimensional case

[4] An equitable rectangular dissection of an $n \times n$ lattice region is a partition into rectangles all of which have area n, where $n = 2^k$ for an even integer k.

[5] A dyadic tiling is a rectangular dissection where each rectangle is dyadic which means that is has the form

$$R = [a2^s, (a+1)2^s] \times \left[b2^t, (b+1)2^t\right],$$

where s, t, a, b are non-negative integers. A dyadic tiling of a $2^k \times 2^k$ square is then a set 2^k dyadic rectangles each of area 2^k whose union covers the square.

[6] A triangulation of a rectangle on $\mathbb{Z}^2$ is a maximal set of non-crossing edges (which are straight lines) each of which connect to lattice points of the rectangle and does not pass through any other point.

for sufficiently large regions. The major part of the chapter is dedicated to random generation. We lay down all the settings such as the height function, flips and connectivity of the configuration space as they appear in [82]. We define a Markov chain MC_{square} for square tilings that is ergodic due to flip-accessibility of the tiling space. Our main interest is to bound its mixing time or, in other words, the time it takes for the chain to reach stationarity. We show fast mixing for long thin regions in Theorem **2.22**. We also consider a weighted version of MC_{square} in which a weight λ is put on big squares. We show a relation with independent sets and in Theorems **2.25** and **2.26** present polynomial upper bounds on the mixing time for MC_{square} with a condition on λ. Finally, we present simulations and conjecture that the mixing time is polynomial for tilings by 1×1 (small) and 2×2 squares but is sub-exponential when $s > 1$. We also show simulations for tilings of a diamond-shaped region that seem to have a frozen part.

In Chapter 3 we study tilings on the Kagome lattice using results from [6]. In the mentioned work the height function was introduced, a fast algorithm for tileability of a given region was shown which also prove the connectivity of the configurational space via local changes (in the same way as for dimer and square tilings). We follow a similar logical path as in Chapter 2. We show how to do approximate counting on a torus using the Kagome tilings' connection to the set of perfect matchings of a corresponding 6-regular graph. We briefly go into the questions of counting perfect matchings. We introduce a Markov chain on the set of Kagome tilings. The connectivity of the tiling space gives us ergodicity of the studied chain. As it was already pointed out, the connectivity of the tiling space is a prerequisite for a Markov chain study. As always we are interested in analyzing its convergence rate. We show why it is not as easy as might seem and propose two alternatives. The first one is to consider a weighted version of the chain and put more weight on one particular tile than on the others. As for square tilings, under a certain condition on the weight λ the weighted chain becomes fast mixing and we draw a polynomial bound in Theorem **3.4**. The second alternative is to consider Kagome tilings using only two types of tiles (lozenges and trapezes). We refer to these tilings as restrained Kagome tilings. We show that for a lozenge-shaped region the tiling space stays connected under local transformations that include only the two types of tiles. In Thereom **3.7** we show that the Markov chain on the restrained Kagome tilings takes no more than $O(N^4)$ steps to mix (N being the area of the region). We present some simulations that show coupling in $O(N^{2.5})$ steps suggesting that our bound is not optimal. At the end of the chapter we provide simulations of a lozenge-shaped region with a linear boundary that seems to have a phenomenon similar to the Arctic circle for domino and lozenge tilings.

Chapter 1

Dimer tilings

1.1 Introduction

The **dimer model** is the study of the set of **perfect matchings** of a graph. A perfect matching of a graph is an independent edge set in which every vertex of the graph belongs to exactly one edge of the graph. The most well-known examples are dimer models on square and triangular grids of $\mathbb{Z}^2$. Dimers on the square grid are equivalent via duality to **domino tilings** that are tilings by 1×2 and 2×1 rectangles. Figure 1.1 shows a domino tiling on a chessboard-coloured region and the corresponding perfect matching of a bipartite graph. Figure 1.2 shows an example of a domino tilings, where four colors correspond to four different tiles: there are two horizontal and two vertical tiles on the chessboard depending on the placement of the white and black units. On the triangular grid dimers correspond to **lozenge tilings**, which are tilings by regular lozenges with a 60 degree rotation (see Figure 1.4)

In Section 1.2 we consider the main results about the counting problem for dimers. In 1961 Kasteleyn [52] calculated the number of domino tilings of a square region using adjacency matrices. Kenyon in [56] showed how to use the inverse Kastelyen matrix to

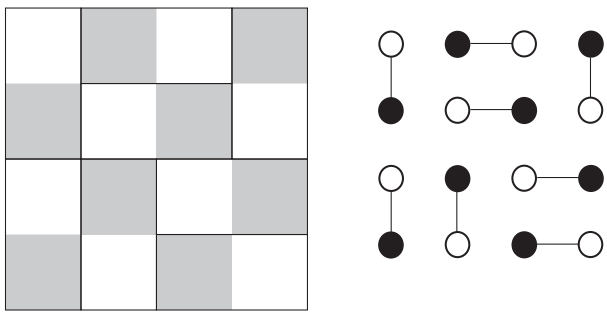

FIGURE 1.1: Domino tiling of a 4×4 square and the associated perfect matching.

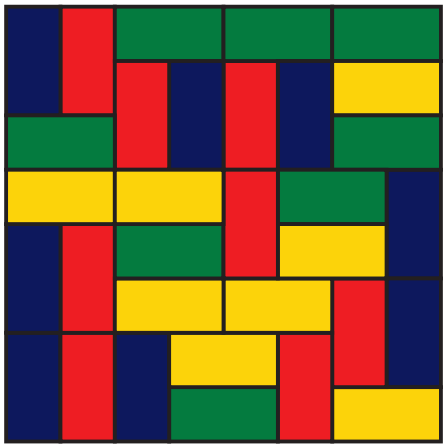

FIGURE 1.2: Domino tiling of an 8×8 square.

calculate the probabilities of particular edges appearing in a uniformly chosen matching. Some 30 year later after Kasteleyn, Elkies et al. [26] considered tilings by dominoes of a so called **Aztec diamond** (see Subsection 1.2.3) and calculated the limit of the **entropy** which is a natural logarithm of the number of tilings divided by the size of the region. They have noticed that random tilings of the Aztec diamond have a peculiar shape with the increase of the size. The **Artic circle** phenomenon appeared. In 2001 Cohn, Kenyon and Propp [16] formulated a variational principle, that among other things, helped the understanding of why random tilings of a square region look homogeneous while those of the Aztec diamond do not at all. They showed that for any simply-connected region the entropy has a limiting value as a function of the shape of the boundary of the region when its size goes to infinity which is unique and corresponds to an asymptotic height function of the region.

One of the problems in the study of dimer tilings is to generate them uniformly at random. Even though approximate counting – which is one of the main reasons that sampling is used for – is not of great interest since there exist formulas for exact counting, the random generation is an interesting problem on its own and is discussed in Section 3.3. The problem of random sampling from uniform distribution was successfully studied by Luby, Randall, Sinclair in [66]. By introducing an appropriate Markov chain and using coupling arguments they proved that the dynamics on dimers is rapidly mixing (in polynomial time over the number of tiles) and gave bounds on the mixing time, which were later improved by Wilson[99], Caputo, Martinelli, Toninelli [14] and Laslier, Toninelli [62]. Wilson and Propp also developed the **Coupling From The Past** algorithm for sampling from stationary distribution (see Appendix B).

Since the studies of domino and lozenge tilings are very related, we are going to cover known results and properties about them at the same time. According remarks will be made about similarities and differences concerning the study of the two models.

1.2 Counting dimers

The problem of counting dimers attracted a lot of attention and there are different ways to calculate the number of dimers of a finite tileable region of $\mathbb{Z}^2$.

The counting problem also depends on the region one considers. The first and classical result on counting dimers [52] was for the domino tilings of a rectangular region of $\mathbb{Z}^2$. A more sophisticated region (the Aztec diamond) was considered by [26] that led to different methods and bijections. Tilings of this region turned out to have interesting long range properties. Analogous region for the lozenge tilings is a regular hexagon, that itself is the most natural region to look when dealing with lozenge tilings. And it turned out that lozenge tilings of a hexagon also possess long range properties similar to the domino tilings.

1.2.1 Adjacency matrices

In 1961 Kasteleyn [52] and Temperley and Fisher [93] calculated independently the number of domino tilings of a rectangular region of $\mathbb{Z}^2$.

Theorem 1.1 (Kasteleyn)**.** *For a closed region R of $\mathbb{Z}^2$ the number of domino tilings equals $\sqrt{|det\ K|}$ where K is the adjacency matrix of the dual graph G with horizontal edges weighted by 1 and vertical edges weighted by $i = \sqrt{-1}$.*

For example, for a 2×3 region the adjacency matrix K is of size 6×6 and $|det\ K| = 9$. For a rectangular region of size $m \times n$ and the associated $m \times n$ bipartite graph G with vertices $\{1, 2, \dots m\} \times \{1, 2, \dots n\}$ the determinant of its adjacency matrix can be calculated using their eigenvalues [57].

It was be shown that for $j, k \in [1, \dots, m] \times [1, \dots, n]$

$$f(x, y) = \sin \frac{\pi j x}{m+1} \sin \frac{\pi k y}{n+1}$$

is an eigenvector of K corresponding to the eigenvalue

$$2 \cos \frac{\pi j}{m+1} + 2i \cos \frac{\pi k}{n+1}.$$

The functions f form an ortogonal basis of eigenvectors and the determinant of K can be written as a product of eigenvalues:

$$\det\ K = \prod_{j=1}^{m} \prod_{k=1}^{n} 2 \cos \frac{\pi j}{m+1} + 2i \cos \frac{\pi k}{n+1}.$$

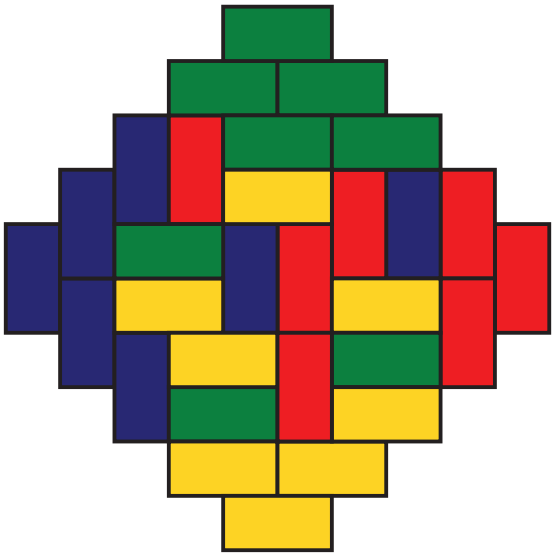

FIGURE 1.3: Domino tiling of the Aztec diamond of order 5.

1.2.2 Plane partitions

Counting the number of the lozenge tilings of a center-symmetric hexagon turned out to be related to the plane partitions if one imagines a lozenge tiling as piles of cubes in a box. This allows to count the number of lozenge tilings using [68]:

$$\prod_{i=1}^{r}\prod_{j=1}^{m}\frac{c+i+j-1}{i+j-2},$$

where r, c, m are the sides of the hexagon. For $n = 20$ this number is approximately 1.6×10^{136}.

1.2.3 Aztec diamond

The Aztec diamond of order n is the union of lattice squares of size 1 that lie inside the area $\{(x, y) : |x| + |y| \leq n + 1\}$. It is therefore composed of $2n(n + 1)$ unit squares. For the Aztec diamond of order 1 there are two possible tilings: placed vertically or horizontally. See Figure 1.3 for an example of a tiling of the Aztec diamond of order 5.

The Aztec diamonds were introduced and studied in [26]. The main result is that the number of domino tilings of the Aztec diamond of order n equals $2^{n(n+1)/2}$. The result is actually broader and proves a general formula for domino tilings of the Aztec diamond. Namely, let T be a tiling of the Aztec diamond of order n, $v(T)-$ the half of the number of vertical tiles in T, $r(T)-$ the minimal number of local moves (**flips**) required to get from the tiling with all horizontal tiles to T. A flip is a change of two vertical dominoes to two horizontal in a square of size 2 and vice versa (it is defined again in section 3.3; for

illustration see the lower part of Figure 1.5). Define the generating function $AD(n; x, q)$ on the set of tilings of the Aztec diamond of order n as follows:

$$AD(n; x, q) = \sum_T x^{v(T)} q^{r(T)},$$

where T varies over all tilings of the order-n Aztec diamond. Then the main result of [26] states that

$$AD(n; x, q) = \prod_{k=0}^{n-1} (1 + xq^{2k+1})^{n-k}. \tag{1.1}$$

When x and q are set to be equal to 1 in (1.1), one gets the number of domino tilings:

$$AD(n) = 2^{n(n+1)/2}.$$

This formula is proven in four different ways: by establishing a bijection between domino tilings and pairs of sign-alternating matrices, finding connection with monotone triangles related to sign-alternating matrices, using the representation theory of general linear groups and finally using the **Domino Shuffling** algorithm [78] which is a more combinatorial approach that allows to get a recurrent formula for the generating function $AD(n; x, q)$. Domino shuffling can also be used to sample domino tilings of the Aztec diamond (see Subsection 1.3.5 for more details).

1.2.4 Inverse Kasteleyn matrix

Using Theorem 1.1, Kenyon [56, 57] obtained a result on the way to calculate the probability of having a certain dimer or a group of dimer appear in a randomly chosen matching of a coloured graph that is dual to a chessboard.

Theorem 1.2. *Let $T = \{t_1, \dots t_k\}$ be a subset of dimers, where $t_i = (w_i, b_i)$ is a dimer covering a white vertex w_i and a black vertex b_i, K is the Kasteleyn matrix of the graph. Then the probability that all dimers in T appear in a u.a.r. matching is*

$$P(T) = |\det(K^{-1}(w_i, b_j))_{1 \le i,j \le k}|.$$

According to this theorem, in order to find a probability of a single edge appearing in a random matching, one needs to look at a single element of K^{-1}. Being able to examine the asymptotics of K^{-1} allows to obtain information about the model's limit behaviour that is discussed further.

1.2.5 Entropy

Topological entropy of the system measures the exponential growth of the number of configurations. Classically, one defines the configurational entropy per dimer ξ for the region tiled with N dimers as follows: let $A(N)$ be the number of dimer tilings, then

$$\xi = \lim_{N\to\infty} \frac{\log A(N)}{N}.$$

One can pose a question of calculating the entropy. For dimer tilings it can be done explicitly since there are closed counting formulas for certain forms of regions. For domino tilings the calculations in [52, 93] allowed to computed entropy for an $m \times n$ rectangular region and for an $m \times n$ torus. ξ was shown to converge to $2G/\pi \approx 0.58$, (where $N = \frac{mn}{2}$) and $G = 1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \dots$ is the Catalan's constant. For the Aztec diamond, on the other hand, it was shown in [26] that the entropy converges to $\log 2/2 \approx 0.338$.

For dimers on the hexagonal lattice for regions the leading term in the entropy depends on the boundary. It was done in [27] using plane partitions. For rhombus tilings of an hexagonal region (that can be seen as a box filled with little cubes, see Figure 1.4) it equals $\frac{3}{2}\log 3 - 2\log 2 \approx 0.261$. The structure of rhombus tilings depends on the boundary and this inhomogeneity is the reason of a difference of entropy between tiling of regions with fixed boundary conditions and free boundary conditions. **Free boundary conditions** mean that the tiles of the tiling are allowed to push out of the boundary of the region. There is a formal relation between the two entropies shown in [19] and in the case of rhombus tilings of a region with free boundary conditions the entropy is larger and is about 0.323.

It is intuitively clear that the entropy value depends on the shape of the tiled region. So another question that attracted interest was to understand which regions maximize the entropy and whether the tilings of such domains have a "typical" look – a property that is almost surely statistically right. It turned out that a typical tiling of an arbitrary finite region can be described by a function that maximizes an entropy integral. Cohn, Kenyon and Propp [16] found an exact formula for the limiting value of the topological entropy ξ for domino tilings as an elliptic integral of a specific function of the shape of the boundary of the region. The integral is called the entropy integral and the function is called the **height function** of a tiling and it will be defined in 3.3.1. In two words, it is an integer valued function defined on the vertices of the grid of the tiled region, and there is a one-to-one correspondence between height functions and tilings. Let us point out that there are cases when the formula is much simpler. It is so for the **isoradial** dimer models – dimer models whose underlying graph is **isoradial** (for example, $\mathbb{Z}^2$, the

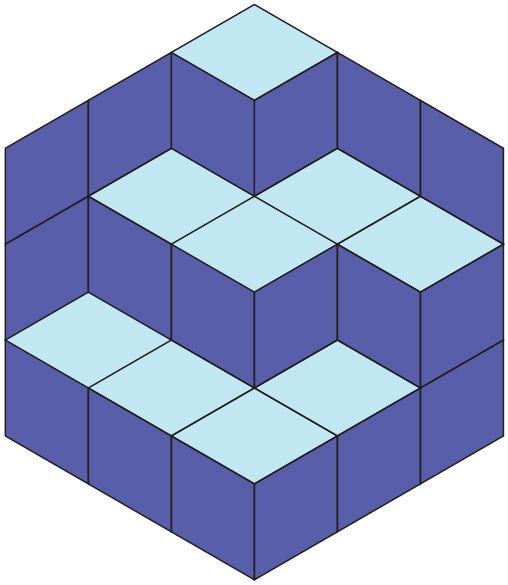

FIGURE 1.4: Lozenge tiling of an hexagon of size 3.

honeycomb and triangular lattices) [1]. For the toroidal partition function it is possible to get an explicit formula for its growth rate that depends only on the geometry of the underlying graphs [95].

1.3 Random generation

Random generation of dimer tilings has attracted attention of a fair number of researchers (for example, [14, 62, 66, 99]). Not only it allows to answer questions from statistical physics about properties of positioning of particles in a dimer model (see, e.g., [38]), it also permits to analyze the limit behaviour of the system. For domino tilings of the Aztec diamond regions [15, 86] and for lozenge tilings of hexagonal regions [17] the long-range order occurs – what happens in the center of the region depends on the boundary. This phenomenon is often referred to as **the Arctic circle** and will be discussed in 3.3.7.

Being able to randomly generate a dimer tiling is useful for studying properties of dimer systems. Note that a fast algorithm for random sampling permits to do approximate counting, but it is of less interest since there are explicit ways of counting that were described in Section 1.2. Let us point out that this will not be the case in Chapters 2 and 3 where square and Kagome tilings will be considered.

Markov chains are useful tools for random generation when there is a finite set of configurations from which one wishes to sample. The idea is to have a connected state space

[1] A graph is called isoradial if is can be embedded in the plane such that every face is inscribed in a circle of radius 1 and all circumcenters of the faces are in the closure of the faces.

Ω; then the Markov chains starts from an initial configuration from Ω and performs a random walk on the graph moving from one configuration to another. After a sufficiently large amount of steps, the Markov chain will converge to a chosen distribution over the entire set of configuration. One can think of it as taking a brand-new decks of cards that is ordered and starting to do random shuffles until the deck is well mixed.

The initial state can be any preferable tiling of the considered region – for example, a tiling made of only horizontal dominoes for a domino tiling; for a lozenge tiling of a right hexagon it can be a tiling that looks like an empty box if one looks at it as a three-dimensional picture. At each step the algorithm chooses a site of the region and performs a local move called a **flip**. For domino tilings a flip is transformation of a 2×2 block: from two horizontal tiles to two vertical tiles and the other way around. For lozenge tilings a flip is a 180 degree rotation of an hexagon of side 1 tiled with three lozenges (see Figure 1.5). In terms of a three dimensional visualization, for lozenge tilings a flip is simply placing or taking away a little cube of size 1 in the big box. For domino tilings there is a similar interpretation where a flip can be seen as moving a hexagonal figure that corresponds to two dominoes (see [94]). When talking about tilings, we often use the term **tiling graph** – it is a graph associated with the set of tilings of a given region in which each vertex corresponds to a tiling and two vertices share an edge if one can get from one tiling to the other by performing a flip.

If one performs these local transformations a sufficiently large number of steps, it is intuitively clear that one should get a random tiling if the configuration space is connected and probabilities of flips are symmetric. The fact that the configuration space is connected follows from a result by Thurston [94]. The question is, for how long one should go on making flips (or performing shuffles in the example with a deck of cards)? How can one know when to stop? This is where the general Markov chain mixing theory comes to help (see, e.g., [63, 74, 80]).

1.3.1 Height function

Dimer tilings can be seen as discrete surfaces in dimension $2 + 1$ with the help of the **height function** introduced by Thurston [94] using Conway's tiling groups [18]. It is a integer-valued function defined on the vertices of a region. Let R denote the region. Let us define it for domino tilings (it is defined in a similar way for lozenge tilings). Choose an origin and set its height as 0 (or any other integer). Consider an edge $e = (x, y)$. If e is not crossed by a domino, set the height difference $h(y) - h(x)$ to be equal to 1 if the square on the left from e is black and -1 if it is white. If e is crossed by a domino, $h(y) - h(x) = 3$ if the square on the left from e is black and -3 if the square to the left

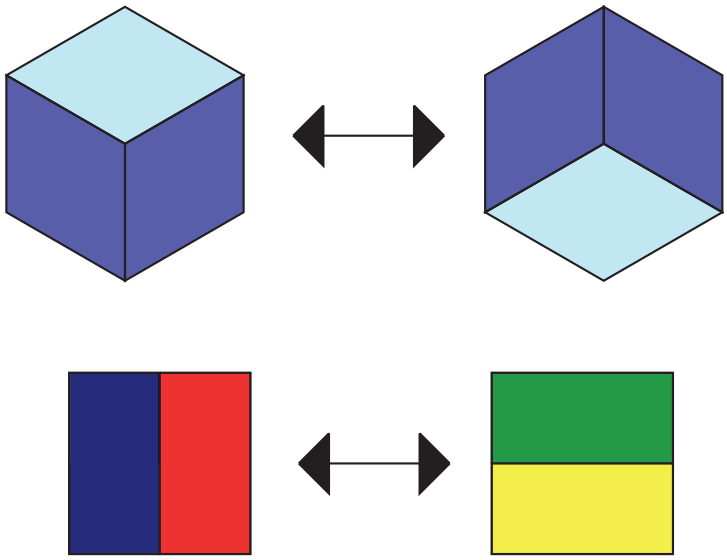

FIGURE 1.5: Single flips for lozenge (top) and domino (bottom) tilings.

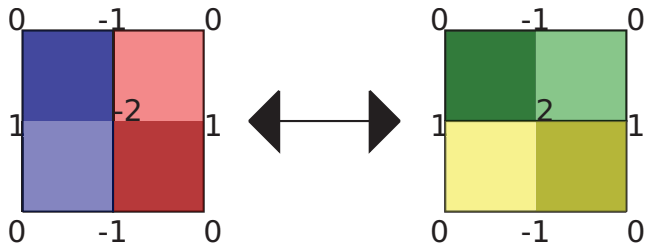

FIGURE 1.6: Single flip for domino tilings with heights.

is white (see Figure 1.6). If one follows a closed oriented cycle and defines the heights in all lattice points of the path and writes down the differences of heights in neighbouring vertices (each of which equals + or -1 depending on the orientation), the total sum of differences will be 0. This construction defines $h(x)$ for all vertices of a given domino tiling that are lattice points of the region up to an additive constant. Note that the height function on the boundary of R does not depend on the tiling. A tiling can be reconstructed in a unique way just be looking at the height function just by starting at a point on the boundary and putting a tile depending on the height of the neighbouring vertices. Each tiling defines a set of heights in each vertex. Two different tilings define two different set of heights. There is a bijection between the set of domino (lozenge) tilings and a class of height functions and random tilings can be seen as random discrete three-dimensional surfaces [58].

A flip around a vertex in a domino tiling changes its height by 4. A vertex x is a **local minimum(maximum)** if $h(x) < h(y)$ ($h(x) > h(y)$) for all neighbours y of x. A flip then changes a local minimum into a local maximum. By performing flips in a tiling that only decrease the height of its vertices, one gets to a **minimal** tiling – tiling in which all local maxima are on the boundary of R. [94] shows a polynomial algorithm

that checks tileability of a simply connected region in $O(n \log n)$, where n is the area of the region. Moreover, it gives an actual tiling as an output and shows that there exists a unique minimal tiling and every tiling can be reduced to the minimal tiling by a finite number of flips which is at most n. This proves connectivity of the set of tilings Ω of R. The height function provides a partial order for the set of tilings: for two tilings A, B of R let h_A, h_B be their height functions, then $A \leq B$ if $h_A(x) \leq h_B(x)$ for all vertices x of R. Then minimal tiling is small than every tiling of the region. The tileability of a simply connected region can actually be checked (much) faster – in $O(p \log p)$, where p is the perimeter, – as states the recent result by Pak,Sheffer,Tassy [76].

1.3.2 Classical Markov chain $\mathbf{MC}_{simple}$

In order to generate dimer tilings uniformly at random using the simple algorithm described above, on can introduce a Markov chain describing the changes in the tiling and bound the time it takes for the chain to reach stationarity. First of all, this means that the chain has to be ergodic. Moreover, it is useful to have a symmetric probability matrix of the chain, because then one obtains a uniformly distributed tiling.

The most natural Markov chain for dimers is the following one. Let R be the tileable region by dimers, N be the number of inner vertices of R, Ω_R be the set of all tilings of R. We say that a flip in x has two **directions**: *up* if it increases $h(x)$, *down* if it decreases $h(x)$.

$\mathbf{MC}_{simple}$:

in Let $T_0 \in \Omega_R$ be an initial configuration. At each time t:

- choose an inner vertex of R with probability $\frac{1}{N}$,
- choose a direction with probability $\frac{1}{2}$,
- perform a flip in the chosen direction in the tiling T_t if possible thus defining the tiling T_{t+1}, otherwise stay still.

Proposition 1.3. *MC_{simple} has uniform stationary distribution.*

Proof. This MC is irreducible because the probability to reach every tiling $T \in \Omega_R$ is positive. The chain is aperiodic since $P(T, T) > 0$ for every $T \in \Omega_R$. It follows that the MC is ergodic and there exists a unique stationary distribution. Moreover, the probability matrix of this chain is symmetric: indeed, $P(T, S) = P(S, T)$ for all pairs

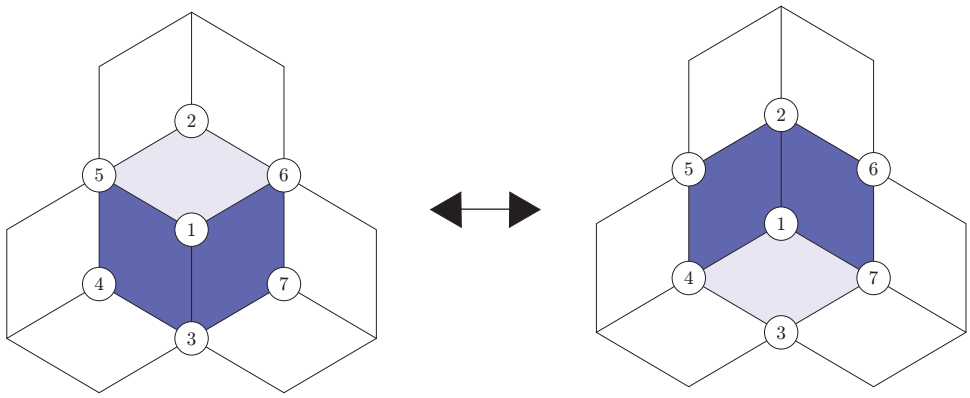

FIGURE 1.7: Two local configurations that are different by one flip and get further from one another in expectation.

(T, S) from Ω_R that are different by one flip. Therefore, the stationary distribution is uniform. □

Since the single-flip Markov chain is monotone (because of the partial order defined by the height function), one can run the **Coupling from the Past** algorithm (see Appendix B) to get a sample from the uniform distribution.

In general, it is a challenging question to bound the mixing time. Usually, to prove that a chain is fast mixing, one uses such methods as the path coupling (see Appendix A) or canonical paths [89]. These methods do not seem to work for MC_{simple}. To give an idea why the problem is not as easy as it might seem, let us show why the path coupling lemma ([10] and Appendix A), which is a handy method to prove that a chain fast mixing, fails.

Why classic path coupling does not seem to work for MC_{simple}:

In order to use path coupling, one has to consider a coupling of a chain and two configurations that are different by one flip and show that the distance between them (generally taken as the Hamming distance on the corresponding tiling graph) does not increase in average at the next step of evolution. More formally, one need to show that if $(T, Q)_t$ is the coupling and $\varphi(T_t, Q_t) = 1$, then $\mathbb{E}(\Delta\varphi_t) \leq 0$, where $\Delta\varphi_t = \varphi(T_{t+1}, Q_{t+1}) - \varphi(T_t, Q_t)$. Figure 1.7 shows two configurations for which the condition above is not satisfied. If the vertex chosen to flip is 1, then flips in both directions decrease the distance: $-\frac{1}{2N} + (-\frac{1}{2N}) = -\frac{1}{N}$. In each of vertices 2,4,7 a flip can be performed only in one direction and it increases the distance: $+\frac{1}{2N}$. All other vertices don't affect the distance change, so $\mathbb{E}(\Delta\varphi_t) = 0$. This gives in total: $\mathbb{E}(\Delta\varphi_t) = +\frac{1}{2N} > 0$.

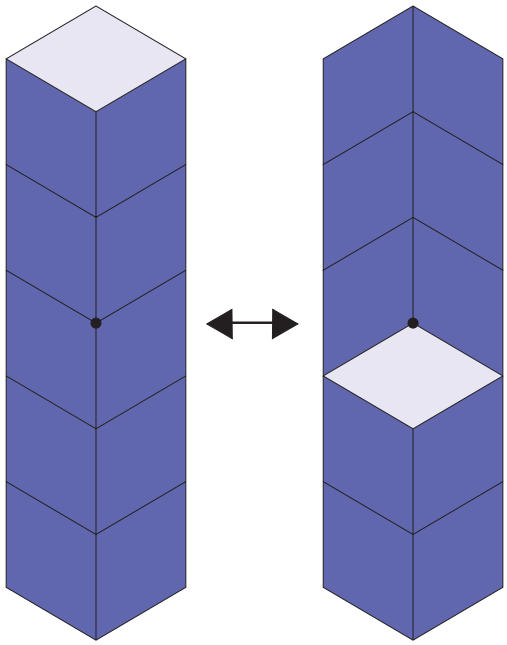

FIGURE 1.8: Tower of flips of height 3 for a lozenge tiling.

There has been a lot of research [14, 66, 99] going on in this direction and researchers working on this problem have used various techniques to get polynomial bounds on the mixing time and then tighten them that are briefly mentioned below.

1.3.3 Modified Markov chain $\mathbf{MC}_{tower}$

In [66] Luby, Randall and Sinclair considered a more complicated dynamics for domino and lozenge tilings which turned out to be easier to analyze. This dynamics is based on the **tower of flips**. In the case of lozenge tilings, a tower is a vertical pile of cubes stuck together. A flip of a tower then is a simultaneous flip up or down of the whole starting from the chosen vertex (see Figure 1.8).

$\mathbf{MC}_{tower}$:

Let $T_0 \in \Omega$ be an initial configuration. At each time t:

- choose an inner vertex of R with probability $\frac{1}{N}$,
- choose a direction with probability $\frac{1}{2}$,
- if the vertex defines a tower of height h, perform a flip in the tiling T_t in the chosen direction with probability $\frac{1}{h}$ thus defining the tiling T_{t+1}, otherwise stay still.

Proposition 1.4. *MC_{tower} has uniform stationary distribution.*

Proof. Note that in the case when $h = 1$, the chain just performs a simple flip. So MC_{tower} is irreducible and aperiodic. And the probability matrix is always symmetric. □

The path coupling lemma now work perfectly fine and gives a polynomial bound on the coupling time.

Theorem 1.5 ([66]). *Let R be a tileable simply connected region of area N. Then MC_{tower} is rapidly mixing and τ_{mix} satisfies*

$$\tau_{mix}(\varepsilon) \leq 8eN^4\lceil \ln \varepsilon^{-1} \rceil$$

Idea of the proof: The path coupling argument is applied to non-intersecting lattice paths (*routings*) which are in bijection with with dimer tilings.

Combining Fourier analysis with path coupling method, Wilson [99] improved the bounds on the mixing time for the lozenge tilings.

Theorem 1.6 ([99]). *Let R be a regular hexagon of area N. Then MC_{tower} is rapidly mixing and τ_{mix} satisfies*

$$\tau_{mix}(\varepsilon) = \Theta(N^2 \log N \lceil \ln \varepsilon^{-1} \rceil)$$

Idea of the proof: The path coupling argument is applied to non-intersecting lattice paths, using a non-integer *displacement* function to measure the distance between configurations instead of the usual flip distance.

Note: The theorem works for any region, not only for simply connected ones. Below is the mixing time upper bound in general case.

Theorem 1.7 (Theorem 7 [99]). *For the tower-flip Markov chain on lozenges on a region which has width w, has m (local) moves separating the top and bottom configurations, and contains p places where a lattice path* [2] *may be moved, after*

$$\frac{2 + o(1)}{\pi^2} p w^2 \log m\varepsilon^{-1}$$

steps, coalescence occurs except with probability ε. For regions which are not simply connected, the coalescence is meant within a given connected component of the state space.

[2] In the previous theorem $p \leq n$ is just the length of each lattice path associated with tilings

1.3.4 Back to single flips

Since the Markov chain with the tower of flips is rapidly mixing, it is natural to wonder whether it could help to analyze the initial chain. It turns out that since MC_{tower} is decomposable into simple flips, one can use the Markov chain decomposition theorem [81] which compares the spectral gaps of the two chains using methods by Diaconis and Saloff-Coste [21] and derives a bound on τ_{mix} for MC_{simple}.

Theorem 1.8 (Theorem 5 [81])**.** *Let R be a tileable simply connected region of area N. Then MC_{simple} is rapidly mixing and τ_{mix} satisfies*

$$\tau_{mix}(\varepsilon) = O(N^4 \log N + N^3 \log N \log \varepsilon^{-1})$$

Idea of the proof: By bounding the length of paths defined between pairs of configuration and plugging it into the decomposition formula from Theorem 3 [81].

Another way of bounding the mixing time of MC_{simple} is to decompose the evolution of the chain into several stages and draw bounds for every stage. Then by linearity of expectation, the mixing time will be bounded by the sum of expected times that the chain spends in every stage and the mixing time for the last stage. This approach was successfully applied by Caputo, Martinelli and Toninelli in [14] and yields a better bound on MC_{simple} in the case of regions with a **good planar boundary** condition. It means that the distance between the boundary and the discrete surface is bounded by a logarithmic factor (see section 2.3 in [14] for full definitions).

Theorem 1.9 (Theorem 2 [14])**.** *For a region with a good planar boundary condition with slope $\boldsymbol{n}$ with a diameter N, for N sufficiently large, MC_{simple} is rapidly mixing and*

$$\tau_{mix}(\varepsilon) \leq N^2 (\log N)^{12} \log \varepsilon^{-1}$$

Remark 1: Under such initial boundary condition, the boundary at equilibrium is essentially flat which is the main ingredient used in the proof of Theorem 1.9.

Remark 2: It was later improved by Laslier and Toninelli in [62].

1.3.5 Domino shuffling

As it was already mentioned, the domino shuffling process can be used to count the number of domino tilings of the n-order Aztec diamond [26]. It can also be used to generate them uniformly at random. For detailed description and proofs see [78]. The

general idea is to go from the n-order diamond to the $n+1$-order diamond by deleting some "bad" blocks of size 2×2, moving the rest a bit which moves the boundary to the size $n+1$, and complete the obtained partial tiling with blocks of size 2 randomly choosing out of two possibilities of tilings for each block. This gives a recurrent formula for the number of tilings and allows to generate them uniformly at random. A variation of the domino shuffling on Aztec half-diamonds was proposed in [73].

1.3.6 Limit shape

One of the best-known results on domino tilings is the Arctic circle theorem in [49]. It describes the asymptotic behaviour of uniformly random domino tilings of the Aztec diamond. It states that for a sufficiently large n, all but a little fraction of the domino tilings of the Aztec diamond of size n will have 4 frozen regions in the four corners of the diamond with dominoes of the same type (there are 4 types of dominoes) piled in a brickwork pattern. Moreover, there will be the **inscribed circle** – the circle of radius $n/\sqrt{2}$ centered in the middle of the Aztec diamond of order n – inside of which the tiling corresponds in an absolutely random way. The case of lozenge tilings was considered in [17] where they were studied as boxed plan partitions. Figure 1.9 [3] shows a random tiling of an hexagon of size 80. One can easily notice the Arctic circle.

[3] Courtesy of Th. Fernique.

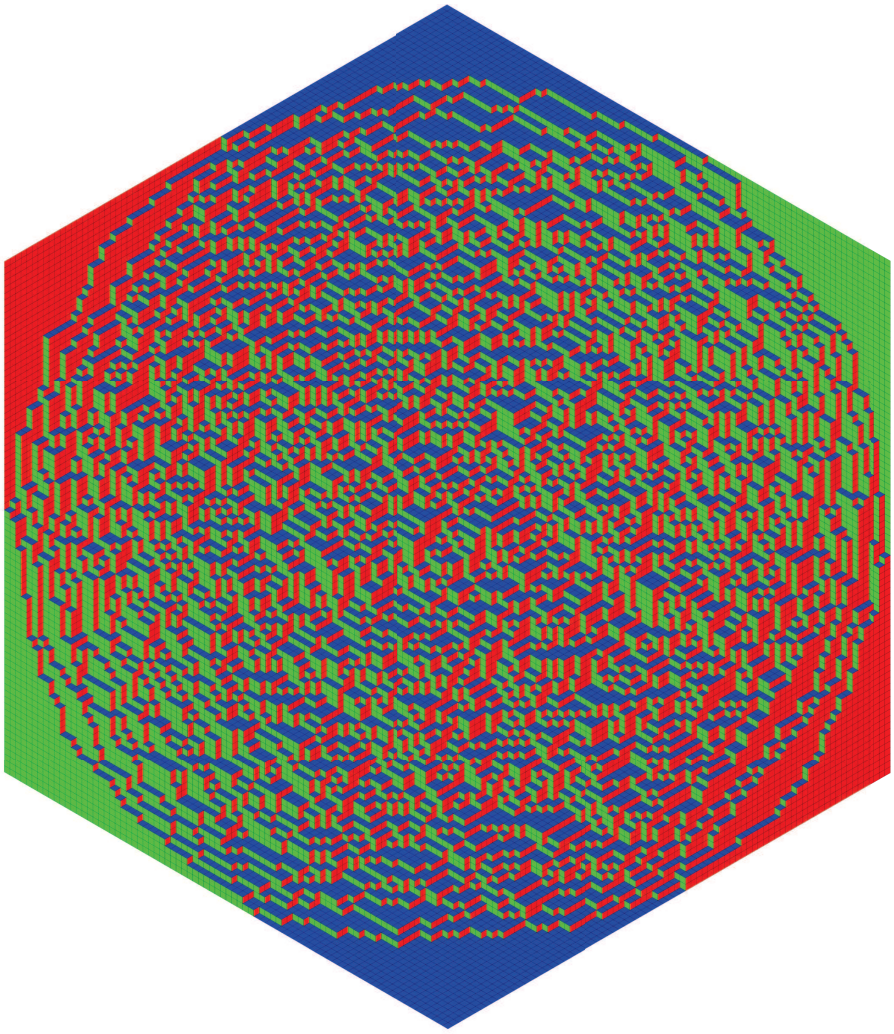

FIGURE 1.9: Random lozenge tilings of hexagon of size 80.

Chapter 2

Square tilings

2.1 Introduction

In this chapter we consider tilings by a pair of squares of different sizes. Denote a tiling by $m \times m$ and $s \times s$ squares where m, s are different positive integers as an $(m, s)-$**square tiling** (or simply a $(m, s)-$ **tiling**). We are essentially interested in the same range of questions which were discussed in Chapter 1 for the dimer tilings. Rectangular tilings seem to be a natural model to consider after dominoes. Here we restrain ourselves to square tilings. As we shall see in throughout this chapter, the study of the square model seems to be already challenging.

First we consider tilings by 1×1 and 2×2 squares. These tilings were studied in [39, 40, 69] from the combinatorial point of view using generating functions and transfer matrices. Considered on a rectangular region, they are in bijection with the *King's problem:* the problem of placing kings on a chessboard in such a way that no king can attack another king. It was studied in [12] using the transfer matrix approach to obtain bounds on the configurational entropy. This problem is also known in statistical physics as a variant of the **Hard square model** and there has been a lot of attempt to find an exact formula for the number of configurations, but with no success.

We use combinatorial tools to understand how the average proportion of small squares in all possible tilings of a $k \times n$ rectangle by small and big squares changes when $k, n \to +\infty$. Another question is to find the number of small squares in tilings of a $k \times n$ rectangle that maximizes the number of tilings of a $k \times n$ rectangle. A simpler problem that we study here is to consider that k is fixed and $n \to +\infty$, which is basically the one-dimensional case. In section 2.3 we propose a generalization of section 2.2.

We pose as well the problem of computing the topological entropy of the square tiling system in Section 2.4. We present a way of approximating the topological entropy using results about subshifts of finite type. Let us point out that the exact solution is only known for a few more tiling models [3, 4, 34, 50, 64] apart from the domino and lozenge tilings that were discussed in the previous section.

In section 3.3, we define height functions for the square tiles as it was done in [54, 55, 82]. We use it to obtain some properties about the structure of the tiling system. In subsection 2.5.3 we introduce a natural Markov chain MC_{square} and show in Theorem 2.23 that for long and thin regions (such as rectangles of size $n \times \log n$) it is rapidly mixing. In subsection 2.5.5 we consider a weighted version of the MC_{square}, where weights λ are put on big squares. We show rapid mixing with certain conditions λ in Theorems 2.26 and 2.27. At the end we provide simulations that raise (a lot of) conjectures for the mixing time for our chains, for the counting problem and for the limit shape of square tilings.

2.2 Combinatorics of $(1, 2)-$square tilings

We will use combinatorial tools to understand how the average proportion of small squares in all possible tilings of a $k \times n$ rectangle by small and big squares changes when $k, n \to +\infty$. A simpler problem that we study here is to consider that k is fixed and $n \to +\infty$.

There has been some work done on the subject. When $k = 2$, tilings of a $2 \times n$ rectangle by 1×1 and 2×2 squares correspond to the Fibonacci sequence. For $k = 3$, one can easily show that the number of ways to cover a $3 \times n$ rectangle with 1×1 and 2×2 squares is equal to $\frac{1}{3}(-1)^n + \frac{1}{3}2^{n+1}$.

Some results were obtained by Heubach [39, 40]. Namely, explicit formulas for the number of tilings for k up to 5 by using introduced basic blocks and methods of analytic combinatorics for finding poles of generating functions and asymptotics. Bigger cases, however, seem to pose problems, mainly because it becomes difficult due to the number of basic blocks.

We use bivariate generating functions (BGFs) to obtain formulas for small cases and calculate distribution of small squares in tilings for $k \leq 10$. We also introduce an automaton construction that represents $BGFs$ and their relations. We extract some properties on its structure, present a simplification algorithm that allows to compute $BGFs$ more easily. We also consider the problem for the cylindrical case. These results were published in [85].

FIGURE 2.1: $4 \times \frac{n}{4}$ board with cut off corners.

2.2.1 Bivariate generating function

In order to study the general case, we introduce $BGFs$. For the sake of simplicity we shall define them for the case $k = 4$ and then generalize the definition. Let

$$Q_{0000}(z, u) = \sum_{n,p} A^4_{n,p} z^n u^p$$

be a BGF where the coefficient $A^4_{n,p}$ of $z^n u^p$ is the number of tilings of a $4 \times \frac{n}{4}$ rectangle with exactly p small squares, supposing that n is a multiple of 4. We want to underline that the rectangle is of area n. This choice is due to the simpler way of defining equations on $BGFs$.

Let $Q_{1000}(z, u)$ be a BGF with the coefficient of $z^n u^p$ being the number of tilings of the initial rectangle with a 1×1 square cut off from the upper left corner and $Q_{2200}(z, u)$ a BGF with the coefficient of $z^n u^p$ corresponding to the number of tilings of the initial rectangle with a 2×2 square cut off from the upper left corner (illustrations are shown in Figure 2.1).

From this point on, we will write $BGFs$ without arguments, always meaning that they are z, u. A relation on Q_{0000}, Q_{1000} and Q_{2200} can be expressed in the following way:

$$Q_{0000} = zuQ_{1000} + z^4 Q_{2200}.$$

Indeed, in order to obtain Q_{0000}, we can either cut off a small square or a big one from the upper left corner of the initial $4 \times \frac{n}{4}$ rectangle. The remaining areas will correspond either to Q_{1000} or to Q_{2200}. And because we cut off squares we need to multiply Q_{1000} by zu (z corresponds to the area occupied by a small square, u – to the one small square) and Q_{2200} by z^4 respectively.

In the same way we can introduce Q_{1100} and Q_{1220} and we have the relation

$$Q_{1000} = zuQ_{1100} + z^4 Q_{1220}.$$

At each step we change indexes of $Q_{i_1 i_2 i_3 i_4}$ by going from left to right in the following way: we permit changing either one 0 to 1 or 00 to 22, which means changing the left one or two columns of the board that was obtained at the previous step by cutting off either a 1×1 or a 2×2 square from the upper left corner of the board. By this rule one can never obtain Q_{1010} or Q_{1022}, for example.

As soon as we get to $Q_{i_1 i_2 i_3 i_4}$ with all indexes being different from zero, we can use a *tetris* rule to reduce the indexes of $Q_{i_1 i_2 i_3 i_4}$ by one layer with "no charge". With "no charge" here means that, given that our strip is infinite, $Q_{1122} = Q_{0011}$, $Q_{1111} = Q_{0000}$ and so on.

Using this technique one obtains a finite set of $BGFs$ $Q_{i_1 i_2 i_3 i_4}$ and a system of functional equations on them. For $k \geq 5$, the principle of constructing a set of $Q_{i_1 \ldots i_k}$ and a system of functional equations is the same.

2.2.2 Counting bounds

In this subsection, for an easier representation, let us denote as $A_l(k)$ the number of configurations of a $k \times l$ rectangle that can be easily obtained from the previous subsection. We are going to show relations and bounds on $A_l(k)$, which are then used to get a way to bound on the entropy (see definition 2.4). Finally we show how using these bounds and dominant singularities of the generating function $Q_{0\ldots0}(z, u)$ to get bounds on the entropy that for an infinite stripe of height $n = kp$, which gets tighter with the increase of k. Table 2.1 contains bounds for $k = 2, \ldots 9$.

Proposition 2.1. *For* $n, k, l \in \mathbb{N}$

$$A_l(n + k - 1) \leq A_l(n)A_l(k) \leq A_l(n + k). \tag{2.1}$$

Proof. The right inequality follows from the fact that when we stick two rectangles of heights k and n together we have a perfect boundary, so the number of configurations is less than in the rectangle of size $k + n$ where there can be big squares on the boundary.

To prove the left inequality let us take a paving of a rectangle of height $k + n - 1$, and look at what happens at height k. On this level we have two types of big squares, those that come from the $(k - 1)$th level and those that come from the $(k + 1)$th level. This means that we can cut the rectangle in two parts, with each part keeping their big squares from the border (kth level), and filling the missing spaces with small squares. The result consists of 2 valid tilings of sizes k and n respectively which can be used to rebuild the initial rectangle of height $k + n - 1$. □

Corollary 2.2.

$$A_l(n) \leq A_l(n-1)\frac{\phi^l - \psi^l}{\sqrt{5}}, \tag{2.2}$$

and

$$A_l(n) \leq \left(\frac{\phi^l - \psi^l}{\sqrt{5}}\right)^{n-1}, \tag{2.3}$$

where $\phi = \frac{\sqrt{5}+1}{2}$ *and* $\psi = \frac{\sqrt{5}-1}{2}$

Proof. Take $k = 2$ and $n-1$ instead of n in the left inequality in Proposition 2.1 and remember that $A_l(2)$ is the $l-th$ Fibonacci number to get (2.2). Then apply (2.2) $n-2$ times to get (2.3). □

Corollary 2.3.

$$A_l(n) \geq \left(\frac{\phi^l - \psi^l}{\sqrt{5}}\right)^{\lfloor \frac{n}{2} \rfloor}.$$

Proof. It suffices to take $k = 2$ in the right inequality in (2.1), apply it n times and use the fact that $A_l(2)$ is the $l-th$ Fibonacci number. □

Let us define the entropy ξ of our system as follows:

Definition 2.4.

$$\xi = \lim_{l,n\to\infty} A_l(n)^{\frac{1}{nl}}.$$

By submultiplicativity that follows from Proposition 2.1, the limit ξ exists. We finally get another corollary from Proposition 2.1.

Corollary 2.5. *Let*

$$A(n) = \lim_{l\to\infty} (A_l(n))^{\frac{1}{l}}, \tag{2.4}$$

then for $n = kp$

$$A(k)^{\frac{1}{k}} \leq \xi \leq A(k+1)^{\frac{1}{k}}. \tag{2.5}$$

Proof. If $n = kp$, then for every $l > 0$ the right inequality of (2.1) gives the following:

$$A_l(k)^p \leq A_l(n).$$

Since $A_l((k+1)^p)$ is a product of $A_l(k+1) \times \ldots \times A_l(k+1)$, one can use the left inequality of (2.1) to get

$$A_l(kp) \leq A_l(k+1)^p.$$

k	lower bound	upper bound
2	1.27201964951407	1.41421356237309
3	1.25992104989487	1.41173913173686
4	1.29513738500437	1.38600758809294
5	1.29841189759956	1.38069947358241
6	1.30842640339979	1.37278796765850
7	1.31203622589229	1.36911693620351
8	1.31639272931298	1.36547044764523
9	1.31901865559441	1.34550944361483

TABLE 2.1: Lower and upper bounds on entropy ξ for $n = kp$, where $k = 2 \dots 9$.

Putting it all together we get

$$A_l(k)^p \leq A_l(n) \leq A_l(k+1)^p.$$

Then

$$A(k)^p \leq A(n) \leq A(k+1)^p.$$

Let us take all parts of the inequalities in the $\frac{1}{n}th$ power:

$$A(k)^{\frac{1}{k}} \leq A(n)^{\frac{1}{n}} \leq A(k+1)^{\frac{1}{k}}. \tag{2.6}$$

If we let p tend to ∞ with k being fixed we get (2.5). □

For $k \leq 9$ bounds on $A(kp)$ can be obtained using dominant singularities $z_{0,k}$ of corresponding GFs using the fact that the coefficients $A_l(k)$ grow like $z_{0,k}^{-lk}$ (the coefficients of the GF have the area parameter). Using Corollary 2.5 one gets the following bound on the entropy:

$$z_{0,k}^{-1} \leq \xi \leq z_{0,k+1}^{-(k+1)/k}.$$

Bounds on the entropy for $n = kp$, where $k = 2, \dots 9$ are represented in Table 2.1. One can notice that these bounds become tighter with the increase of k. The further one could go with calculation, the tighter bounds could be obtained. There is more discussion about topological entropy and mostly about its efficient approximation in Section 2.4.

2.2.3 Combinatorial results

Let us go back to generating functions that were defined in subsection 2.2.1. As it was defined

$$Q_{0000}(z,u) = \sum_{n,p} A^4_{n,p} z^n u^p,$$

where $A^4_{n,p}$ is the number of tilings of $4 \times \frac{n}{4}$ rectangle with p small squares. For the general case:

$$Q_{\underbrace{0\ldots0}_{k}}(z,u) = \sum_{n,p} A^k_{n,p} z^n u^p,$$

where $A^k_{n,p}$ is the number of tilings of $k \times \frac{n}{k}$ rectangle with p small squares.

For each k one gets a system of equations on BGFs where coefficients of Q encode the vertical border of the stripe. Traditional combinatorial tools (see, e.g., [31]) allow us to find formulas for the $BGFs$ and extract some properties. We can solve a system of equations and find $Q_{0\ldots0}(z,u)$ for small k. It starts getting complex for $k \geq 10$ given that the size of the associated matrix grows exponentially.

For example, for $k = 4$, one gets the following solution for $Q_{0000}(z,u)$:

$$Q_{0000}(z,u) = \frac{1 - z^4}{1 - z^4 - z^4u^4 - 2z^8u^4 - z^8 + z^{12}u^4 + z^{12}}$$

$$Q_{0000}(z,1) = \frac{1 - z^4}{1 - 2z^4 - 3z^8 + 2z^{12}}$$

Since u was in charge of the number of small squares, when we set $u = 1$, the coefficients A^k_n of $Q_{0000}(z,1)$ correspond to the number of all tilings of an $4 \times \frac{n}{4}$ rectangles without the count of small squares. These coefficients satisfy the recurrence equation: $a_n = 2a_{n-1} + 3a_{n-2} - 2a_{n-3}$ with $a_0 = a_1 = 1, a_2 = 5$ [A054854] [91].

For $k = 5$:

$$Q_{00000} = \frac{u^2z^{10} + uz^5 - 1}{1 - u^5z^5 - uz^5 - 3u^6z^{10} - 4u^2z^{10} - u^7z^{15} + 3u^3z^{15} - 3u^4z^{20}}.$$

We are interested in calculating the average proportion of space occupied by small squares. The standard technique is to differentiate generating functions and then use the singularity analysis [31]. Let us show step by step how we do this.

For each k the average proportion is represented by the following expression:

$$\frac{\sum_p \frac{n}{p} A^k_{n,p}}{\sum_p A^k_{n,p}}. \tag{2.7}$$

k	z_0	%
2	0.786151377757423	44.7213595499958
3	0.793700525984100	55.5555555555555
4	0.772118859032568	46.9544274320176
5	0.770171624157749	49.5070291158392
6	0.764276842321142	47.2416138330296
7	0.762174077411561	47.7590234411527
8	0.759651719226601	47.0291235954429
9	0.758139390795314	47.0553488489594
10	0.76569872370847	46.7648743206487

TABLE 2.2: $(1,2)-$case: k is the height of the region, z_0 – dominant singularity of the corresponding BGF, % – average percentage of space occupied by small squares.

It can be obtained using the derivatives $\partial_u Q_{0\ldots0}(z,u)$ and $\partial_z Q_{0\ldots0}(z,u)$:

$$\partial_u Q_{0\ldots0}(z,1) = \sum pA^k_{n,p} z^n,$$

$$z\partial_z Q_{0\ldots0}(z,1) = \sum nA^k_{n,p} z^n.$$

So (2.7) is exactly

$$\frac{[z^n]\,\partial_u Q_{0\ldots0}(z,1)}{[z^n]\,z\partial_z Q_{0\ldots0}(z,1)} \tag{2.8}$$

We can use the singularity analysis to estimate the coefficients in front of z^n in (2.8) and get the average proportions of space occupied by small squares. Therefore we calculate the singularities of $Q_{0\ldots0}(z,u)$ that are the closest to zero. The list of singularities for $k = 2,\ldots,10$ is shown in Table 2.2. Denoting the dominant singularity (closest to zero) by z_0, we get

$$\left.\frac{\partial_u Q_{0\ldots0}(z,u)}{z\partial_z Q_{0\ldots0}(z,u)}\right|_{(1,z_0)}.$$

Average proportions of space occupied by small squares for $k \leq 10$ are shown in Table 2.2.

2.2.4 Automaton representation

For each k let us introduce an automaton. Each $Q_{i_1\ldots i_k}$ with $i_j \in \{0,1,2\}$ for $j = 1,\ldots k$ is associated with a state $q = i_1\ldots i_k$ and each functional equation involving $BGFs$ can be translated into an automaton transition. For example, the relation

$$Q_{0000} = zuQ_{1000} + z^4 Q_{2200}.$$

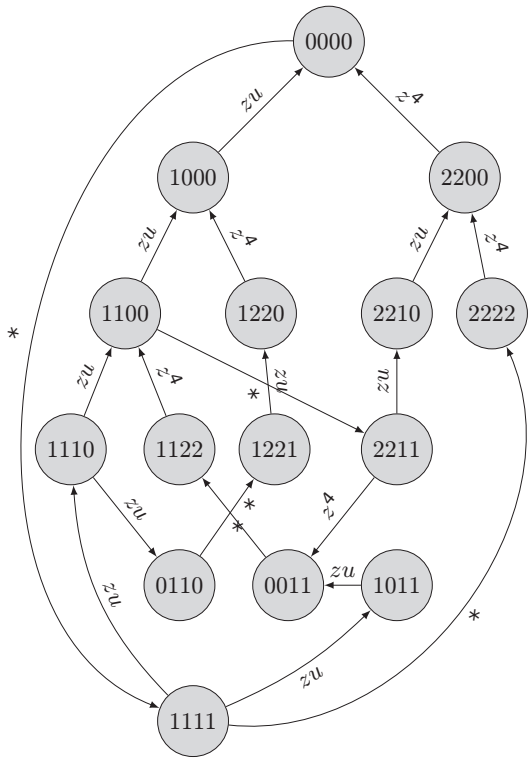

FIGURE 2.2: Automaton for $k = 4$.

is represented in the following way: an arrow marked by zu goes from the state 1000 to the state 0000, an arrow marked by z^4 goes from the state 2200 to the state 0000. When the tetris rule is applied, we will mark the corresponding arrows by *a star*.

Commentary: The unusual way, one might say, of directing arrows can be explained by the fact that adding 1 or 22 to the indexes of $BGFs$ corresponds to cutting off corners of the initial rectangle.

For $k = 4$ the set of states consists of the states: 0000, 1000, 2200, 1100, 1220, 2210, 2222, 1110, 1122, 1221, 1111, 0011, 0110, 1011, 2211 and an illustration of the automaton is shown in Figure 2.2.

We shall refer to the state that consists of all $0s$ as *initial*. Calculation of the paths in the automaton that start and end at the initial state will allow us to find formulas for $Q_{0\ldots0}$. Our objective is to decrease the computational complexity by reducing the number of states, which basically means reducing the number of functional equations in the system.

2.2.4.1 Essential, non-essential and additional states

Definition 2.6. A state q of an automaton is called **essential** if there are at least two arrows coming in and out of q and at least one of the arrows coming out is marked by

a star. It is called non-essential otherwise.

Let E_k be the set of all essential states for each $k \geq 4$. One can see from Figure 2 that $E_4 = \{1100\}$ and $|E_4| = 1$. Let us describe the structure of E_k and find $|E_k|$.

Proposition 2.7. *A state $q = i_1...i_k$ of an automaton is essential if and only if q has the following properties:*

1. *q consists only of $0s$ and $1s$;*
2. *All $1s$ come in consecutive pairs in q;*
3. *$i_1 = i_2 = 1$;*
4. *There are at least two $0s$ and the leftmost 0 in q comes in a pair with another 0.*

Proof. $\Rightarrow$ Let us prove that if q does not have at least one of the four properties, then q is non-essential. If there exists $j \in \{1, ..., k\}$ such that $i_j = 2$ or a block of consecutive $1s$ whose length is odd in q, then no state can be reduced to q by the use of the tetris rule, so no arrow marked by a star comes out of it. If q starts with a zero, then at most one arrow comes out of q (to the state that is reduced to q by the tetris rule). If the leftmost zero in q is isolated, then only one arrow comes in q (from the state with the leftmost 0 being replaced by 1).

$\Leftarrow$ Consider that q has these four properties. Given that the leftmost 0 is not isolated, there are two arrows coming in q. And it is clear that there are two arrows that come out of q – one to the state $q' = i_1 + 1 \dots i_k + 1$ that is reduced to q by the tetris rule and one to a state with a 1 on the left from the leftmost 0 being replaced by a 0.

□

Proposition 2.8. *For $k \geq 4$ the number of essential states $|E_k|$ in the automaton is represented by the following formula:*

$$|E_k| = \{ \begin{array}{ll} F(k\text{-}3) - 1 & \text{if } k = 2l+1 \\ F(k\text{-}3) & \text{if } k = 2l \end{array}$$

where $F(k-3)$ is the $(k-3)rd$ Fibonacci number.

Proof. Looking at essential states of length k is the same as looking at states of length $k-3$ that are obtained from essential states by deleting the first two $1s$ and gluing

together the first two consecutive $0s$ in each essential state. The obtained states are in bijection with tilings of a strip $1 \times k-3$ by blocks of size 1×1 and 1×2 with at least one 1×1 block. □

Corollary 2.9. *With $k \to \infty$*

$$|E_k| \sim \frac{\phi^{k-3}}{\sqrt{5}}$$

Definition 2.10. A state is called **additional** if it belongs to a cycle that does not contain any essential or initial states.

Note that only non-essential states can be additional. The interest of looking at additional states is, merely, because in order to properly reduce an automaton and get the explicit formulas for our generating functions, we need to pay attention to all the cycles in the automaton including the cycles that do not pass through the initial state. If not, we might lose some terms in the resulting formulas.

Our objective is to choose additional states in such a way, so that each cycle in the automaton includes either the initial state, essential states or the chosen additional states. The idea is to minimize the number of additional states that have to be added to the initial and essential states in order to properly reduce the automaton. We are not going to calculate the number of all additional states. Nor will we minimize this number. Rather, we will define a subset of the set of additional states and try to justify this choice by proving that it provides us with the wanted structure.

Let us take a set that consists of states that have the same structure as the essential states but with an odd block of $1s$ of size at least 3 on the left from the leftmost 0. We denote this set by A_k. Clearly,

$$|A_k| = |E_{k-1}|\,.$$

Proposition 2.11. *Each state from A_k is additional.*

Proof. We need to show that every $q \in A_k$ belongs to at least one cycle that does not contain an essential state or the initial state. For $i = 1$ $q_1 = 1110\ldots0$ and the cycle is schematically shown in Figure 3 with $f(z,u)$ and $g(z,u)$ being transition functions between states. There are as many cycles as there are possible ways to get from the state $00011\ldots\ldots$ to the state $111220\ldots0$. It is not difficult to see that there are no essential states between those two states. For other i the structure of cycles that contain q_i is analogous. □

Corollary 2.12. *There are states in A_k for $k \geq 6$ that belong to more than cycle with no essential or initial states contained in it.*

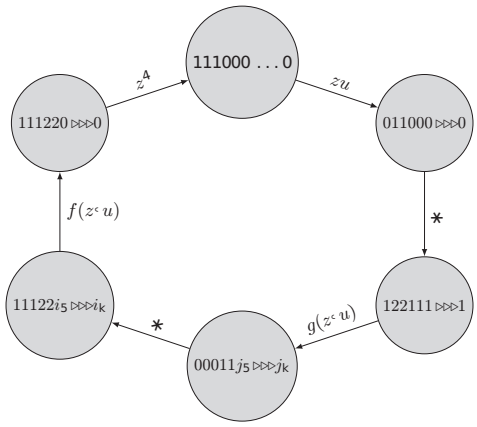

FIGURE 2.3: Cycle with the state $1110\ldots0$ for $k \geq 5$.

Now the question is, if we mark the initial state and all the states from E_k and A_k in the automaton, does it ensure that there are no cycles left that do not contain the marked states? Proving that will justify our choice for keeping these particular states.

Proposition 2.13. *Let $q = i_1 \ldots i_k$ be a state that does not belong to $A_k \cup E_k \cup \{0\ldots0\}$. Then a cycle (or cycles) that q belongs to, contain(s) states from $A_k \cup E_k \cup \{0\ldots0\}$.*

Proof. Let us point out that every state that contains $2s$ belongs to a cycle with a state that contains only $0s$ and $1s$ (the state to which it gets reduced to at some point via the tetris rule). So it is sufficient to prove the statement only for states that consist of $0s$ and $1s$.

Let $q = i_1 \ldots i_k$ with $i_j \in \{0,1\},\ j = 1,\ldots k$ (Figure 2.3 might help to visualize the cycles.) If the leftmost zero in q is isolated, then there is only one arrow coming in this state from a state from $A_k \cup E_k$. If the leftmost zero is not isolated, let $i_1 = \ldots = i_l = 1,\ i_{l+1} = i_{l+2} = 0,\ 0 \leq l \leq k-2$. If $l = 0$, then $q = 00...011i_{n+3}...i_k$ where $i_1, \ldots, i_n = 0$, $n \geq 2$. In this case q belongs to a cycle with the state 11...100...0 belonging to $A_k \cup E_k$. The situation is similar if $l = 1$ apart from the case when $i_j = 0$ for all $k+3 \leq j \leq k$. Then $q = 10\ldots0$ and belongs to a cycle that contains the initial state. For $l \geq 2$ since all $1s$ after the $l+2$ coordinate in q come in pair, q belongs to a cycle (cycles) with a state from $A_k \cup E_k$ where the block with leftmost $0s$ is filled with $1s$. □

Remark: We shall further refer to the states from A_k as **additional***.

2.2.4.2 Simplified automata

We can simplify an automaton by keeping only the initial, essential and additional* states and reducing all other states. The rules of reduction are shown in Figure 2.4. We

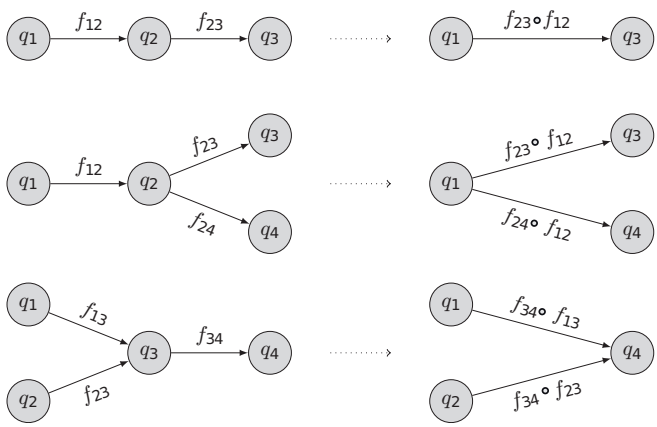

FIGURE 2.4: Rules of reduction for an automaton.

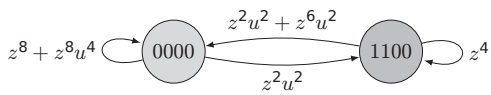

FIGURE 2.5: Reduced automaton for $k = 4$.

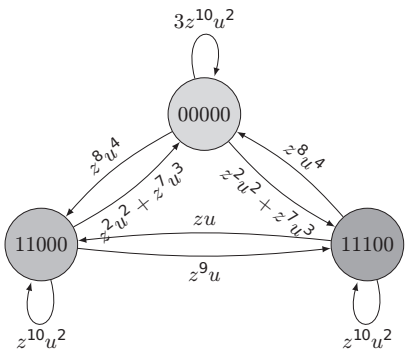

FIGURE 2.6: Reduced automaton for $k = 5$.

denote by f_{ij} a transition between states q_i and q_j which is represented by an arrow going from q_i to q_j.

In the case $k = 4$, there is one essential state 1100 and no additional states. A reduced automaton for the case $k = 4$ is shown in Figure 2.5.

For a $5 \times n$ rectangle there is one essential state 11000 and one additional* state 11100. A reduced automaton is shown in Figure 2.6. For $k = 6, 7$ reduced automata are schematically shown in Figure 2.7.

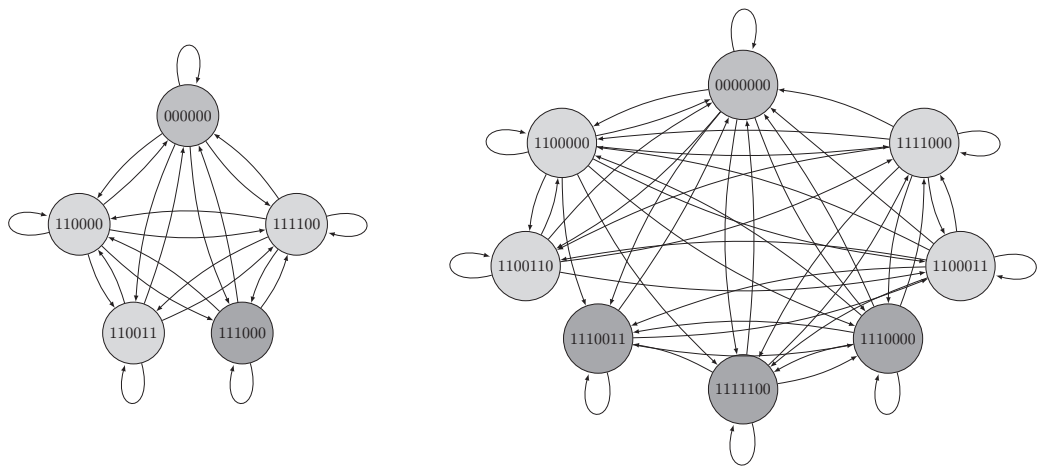

FIGURE 2.7: Reduced automata: for $k = 6$ (left), for $k = 7$ (right).

2.2.5 Matrix representation

Another point of view on this problem comes from the problem of non-attacking kings where one looks at the number of ways kings could be put on a rectangular board without having two king attacking each other. For the sake of completeness we mention this approach and results that were obtained previously.

There is a natural bijection between tilings of a $k \times k$ region by 1×1 and 2×2 squares and configurations of non-attacking kings on a $(k-1) \times (k-1)$ board. One simply has to consider that every big square has a king in his left bottom corner, and having no kings attacking each other is equivalent to having no big squares intersecting.

This problem has been studied using the matrix approach [12]. Namely, the use of the adjacency matrix of the graph where nodes are lines of the rectangle, two nodes are connected if the two lines can be put on top of each other (for the use of this approach see also [69]). These adjacency matrices verify the following recursion:

$$A_0 = (1), A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \ldots, A_n = \begin{pmatrix} A_{n-1} & A_{n-2} \\ A_{n-2} & 0 \end{pmatrix}.$$

Note that the dominant eigenvalue of A_i is equal to the inverse of the singularity of $Q_i(z, 1)$. In [12] it was used to compute good approximation of the entropy but it can be also used to compute the average proportions of space occupied by small squares for a given k. Indeed, for a fixed $k \times n$ region one can compute the number of big squares by taking a matrix M which is a diagonal with the number of big squares in the

corresponding node, and calculate

$$b = \mathbf{1}\left(\sum_{p\le n} A_k^p M A_k^{n-p}\right)\mathbf{1}^T,$$

where $\mathbf{1} = (1,\dots,1)$.Then the proportion of big squares equals

$$\frac{4b}{kc},\ \ where\ \ c = \mathbf{1}\left(\sum_{p\le n} A_k^p I A_k^{n-p}\right)\mathbf{1}^T,$$

where c is the number of lines in all the configurations. We can then obtain result when $n \to \infty$ by putting A_k in Jordan form and dividing both b and c by λ_1^n where λ_1 is the dominant eigenvalue of A_k.

The latest improvement was done by J. Nillson that allowed to compute the exact number of tilings by 1×1 and 2×2 squares of squares of size up to 40 [72], which permitted to extend the sequence [A063443] with 15 new entries.

2.2.6 Cylindrical case

Let us now consider a $k \times \frac{n}{k}$ rectangle with sewn horizontal borders. We obtain a cylindrical region that we want to tile with 1×1 and 2×2 squares. The way of constructing BGFs stays the same with the only difference that now we allow having separated 2s in the first and last positions of the indexes of Generating Functions. The functional equations, therefore, change. For example,

$$Q^c_{0000} = zuQ^c_{1000} + z^4Q^c_{2200} + z^4Q^c_{2002},$$

where the index c is used to distinguish between cylindrical and plain cases.

Let us construct an automaton in the same way as before (see Figure 2.8).

The notions of essential and additional states stay the same. The set E_k^c of essential states in the cylindrical case equals E_k. But the set A_k^c consists not only of states from A_k, but also of E_{k-2} and the initial state of size $k-2$ where $1s$ are added in the first and last positions which might of course create a rightmost block of an odd length but that is due to the fact that there are states with a 2 in the first and the last positions. So $|A_k^c| = |A_k| + |E_{k-2}| + 1$.

In the case $k = 4$ there is one essential state 1100 as in plain case and one additional* state 1001. The rules of reduction stay the same and a reduced automaton is shown in Figure 2.9.

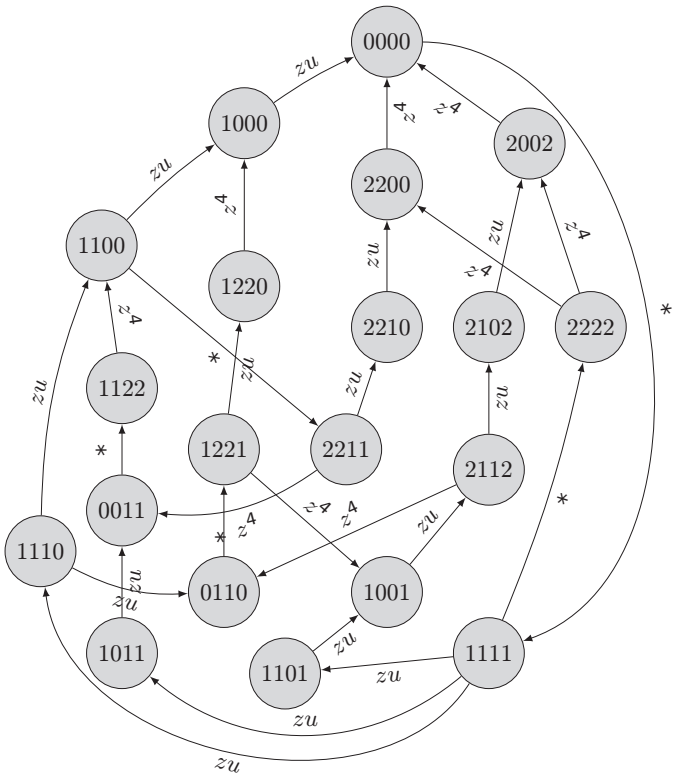

FIGURE 2.8: Automaton for $k = 4$, cylindrical case.

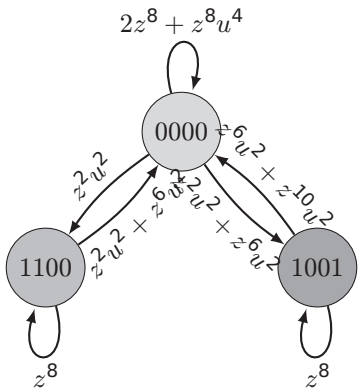

FIGURE 2.9: Reduced automaton for $k = 4$, cylindrical case.

In this case we have

$$Q^c_{0000} = \frac{1 - z^4}{1 - 3u^4z^8 + 2z^12 - u^4z^4 - 2z^8 - z^4},$$

and small squares occupy, in average, 0.466 of the space which is smaller than in the plain case. It is rather understandable – because of the sewn boarders there are less constraints on the way big squares can be placed.

Reduced automata for $k = 5, 6$ are schematically shown in Figure 2.10.

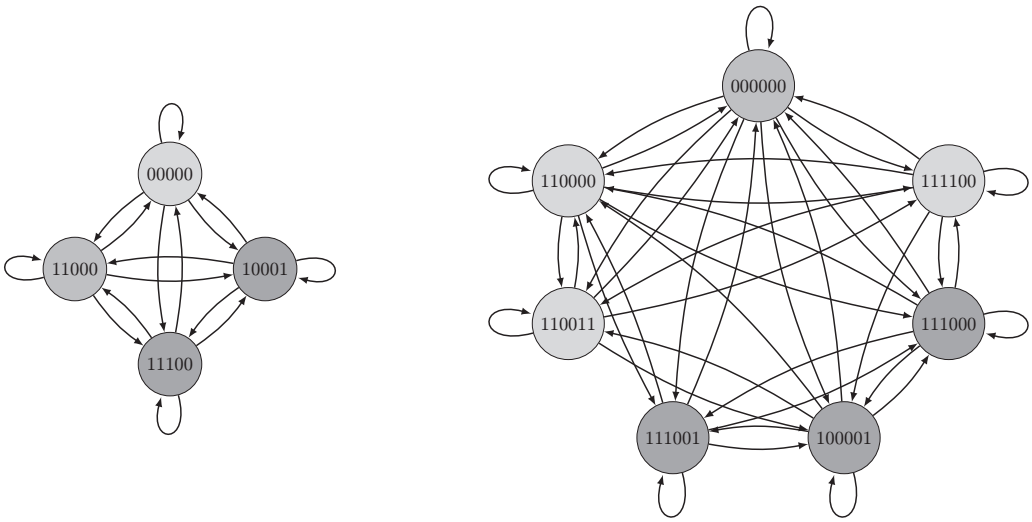

FIGURE 2.10: Reduced automata in cylindrical case: $k = 5$ (left), $k = 6$ (right).

2.2.7 Open problems

The main question that we would like to answer is whether the limit of average proportions for the plain/cylindrical case discussed in this section exists and how to compute it. Calculations that were done by M. Rao using the adjacency matrix approach, give a decreasing sequence from 0.466 to 0.458 from $k = 10$ up to 25. Another question would be to find the number of small squares in tilings of a $k \times n$ rectangle that maximizes the number of tilings of a $k \times n$ rectangle.

2.3 Combinatorics of $(1, s)-$square tilings

The same method using $BGFs$ can be applied to the general case.

Consider, for example, the case of $(1, 3)-$square tilings. Define the $BGFs$ in the same way as for the $(1, 2)-$square tilings. For the case $k = 4$, the equations on the $BGFs$ become the following:

$$Q_{0000} = Q_{3330}z^9 + Q_{1000}uz,$$

and so on. In the same way, on gets the system of the equations.

For example, for $k = 4$:

$$Q_{0000}(z, u) = \frac{-1}{2u^3z^{12} + u^4z^4 - 1},$$

k	z_0	%
3	0.8803	41.723
4	0.8763	49.296
5	0.8829	56.707
6	0.8701	46.084
7	0.8653	45.883
8	0.8659	48.856
9	0.8633	46.784
10	0.8609	45.641

TABLE 2.3: $(1,3)$ case: k is the height of the region, z_0 – dominant singularity of the corresponding BGF, % – average percentage of space occupied by small squares.

k	$\% - 2$	$\% - 3$	$\% - 4$	$\% - 5$
3	55.555	41.723	100	100
4	46.954	49.296	39.665	100
5	49.507	56.707	45.198	38.115
6	47.241	46.084	51.686	42.264
7	47.759	45.883	54.670	41.626
8	47.029	48.856	44.573	43.412
9	47.055	46.784	43.325	44.890
10	46.764	45.641	49.365	39.430

TABLE 2.4: Percentage of average space occupied by small squares for $(1,2),\ldots,(1,5)-$ square tilings of stripes of height $k \leq 10$.

$$Q_{0000}(z,1) = \frac{-1}{2z^{12} + z^4 - 1}.$$

For $k = 5$:

$$Q_{00000}(z,u) = \frac{-1}{3u^6z^{15} + u^5z^5 - 1},$$

$$Q_{00000}(z,1) = \frac{-1}{3z^{15} + z^5 - 1}.$$

Average proportions of space occupied by small squares for $(1,3)-$tilings of stripes of height $k \leq 10$ are shown in Table 2.3 and for $(1,2),\ldots,(1,5)-$ square tilings of stripes of height $k \leq 10$ in Table 2.4.

2.3.1 Comparison

Up until now we were calculating average proportions of space occupied by squares 1×1. Let us denote it by $C(s,k)$ for $(1,s)$ tilings of a $k \times k$ square region. Table 2.5 shows a comparative of proportions $C(s,k)$ of space occupied by small squares for $(1,2)$ and $(1,3)$ cases for squares regions of sizes 10 to 25 plotted in Figure 2.11. Table

k	2: [min] avg [max]	3: [min] avg [max]
10	[0.280] 0.492 [0.639]	[0.280] 0.499 [0.730]
11	[0.372] 0.487 [0.669]	[0.331] 0.515 [0.702]
12	[0.308] 0.486 [0.638]	[0.315] 0.503 [0.682]
13	[0.360] 0.483 [0.621]	[0.360] 0.486 [0.680]
14	[0.387] 0.480 [0.571]	[0.357] 0.481 [0.678]
15	[0.342] 0.479 [0.573]	[0.360] 0.481 [0.639]
16	[0.390] 0.476 [0.593]	[0.367] 0.480 [0.613]
17	[0.348] 0.475 [0.571]	[0.377] 0.472 [0.567]
18	[0.395] 0.471 [0.580]	[0.388] 0.481 [0.593]
19	[0.357] 0.468 [0.560]	[0.351] 0.467 [0.576]
20	[0.370] 0.471 [0.557]	[0.370] 0.468 [0.577]
21	[0.392] 0.469 [0.537]	[0.367] 0.477 [0.550]
22	[0.396] 0.469 [0.537]	[0.386] 0.469 [0.553]
23	[0.379] 0.464 [0.523]	[0.373] 0.478 [0.542]
24	[0.395] 0.467 [0.528]	[0.375] 0.451 [0.531]
25	[0.398] 0.467 [0.532]	[0.366] 0.461 [0.524]

TABLE 2.5: Average space occupied by small squares for $(1,2)$ and $(1,3)-$ square tilings of squares of sizes 10 to 25 over 100 to 1000 trials ($C(s,k)$ with $s = 2,3$ and $k = 10,\ldots 25$.)

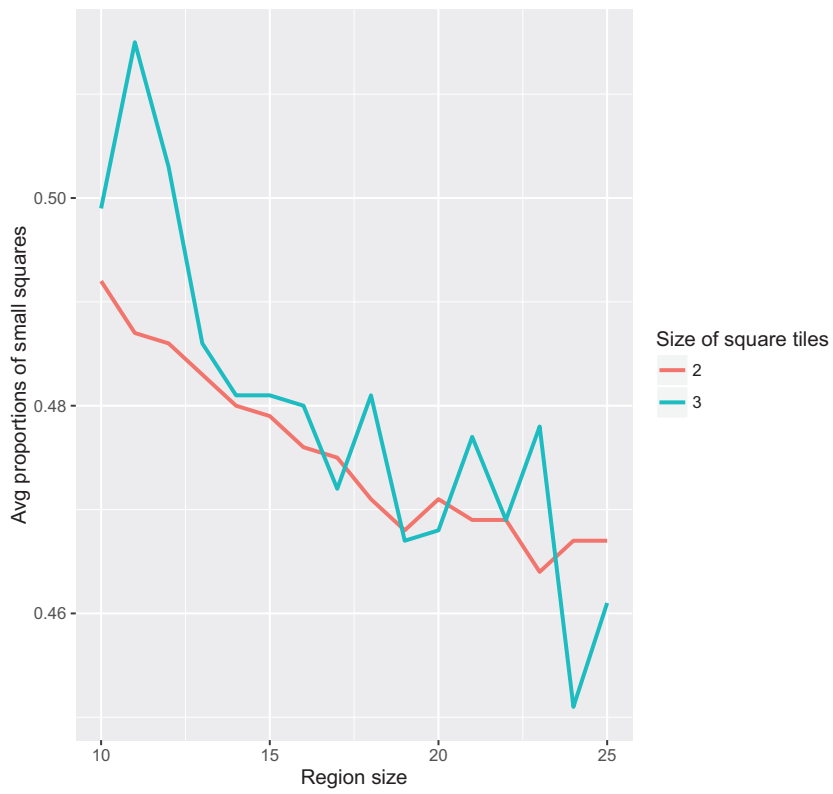

FIGURE 2.11: Plot for average space occupied by small squares from Table 2.5, $x-axis$ represents the size of the tiled square, $y-axis$ marks average proportions.

2.6 shows $C(s,k)$ for $s = 2,\ldots,10$ and $k = 20,25,30$. Let $B_{space}(s,k)$ be the average ratio of space occupied by small squares to space occupied by big squares in tilings of $k \times k$ square. And finally, let $B_{number}(s,k)$ be the average ratio of the number of small squares to the number of big squares. Clearly, $B_{number}(s,k) = s^2 B_{space}(s,k)$ and $B_{space}(s,k) = \frac{C(s,k)}{1-C(s,k)}$. See Table 2.7 for computational results on $B_{number}(s,k)$ for $s = 2,\ldots 10$ and $k = 20,25,30$.

Let

$$C(s) := \lim_{k\to\infty} C(s,k)$$

$$B_{number}(s) := \lim_{k\to\infty} B(s,k)$$

Conjecture. For $(1,s_1)$ and $(1,s_2)$ square tilings with $s_2 > s_1$:

$$C(s_2) \geq C(s_1) - \varepsilon(s_1,s_2). \tag{2.9}$$

Proposition 2.14. *For $(1,s_1)$ and $(1,s_2)$ square tilings with $s_2 > s_1$, if*

$$C(s_2) \geq 1 - \frac{C(s_1)-1}{C(s_1)\alpha - 1},$$

where $\alpha = \frac{s_2^2 - s_1^2}{s_2^2}$, then

$$B_{number}(s_2) \geq B_{number}(s_1).$$

Proof. Consider $\frac{B_{number}(s_2)}{B_{number}(s_1)}$:

$$\frac{B_{number}(s_2)}{B_{number}(s_1)} = \frac{C(s_2)}{C(s_1)}\left(\frac{1-C(s_1)}{1-C(s_2)}\right)\frac{s_2^2}{s_1^2}. \tag{2.10}$$

We are interested in knowing when (2.10) ≥ 1. So we get:

$$(1-C(s_1))C(s_2) \geq (1-C(s_2))C(s_1)\frac{s_1^2}{s_2^2}.$$

Denote $\alpha = \frac{s_2^2 - s_1^2}{s_2^2}$, then $\frac{s_1^2}{s_2^2} = 1-\alpha$. Then

$$C(s_2) \geq C(s_1) - C(s_1)\alpha + C(s_1)C(s_2)\alpha,$$

$$C(s_2)(1-\alpha C(s_1)) \geq C(s_1)(1-\alpha),$$

k	2	3	4	5	6	7	8	9	10
20	0.471	0.468	0.432	0.503	0.45	0.578	0.491	0.582	0.586
25	0.467	0.461	0.436	0.491	0.473	0.444	0.483	0.56	0.481
30	0.462	0.469	0.51	0.472	0.481	0.512	0.521	0.42	0.546

TABLE 2.6: Average space occupied by small squares for $(1,2),\dots,(1,10)-$ square tilings of squares of sizes 20, 25, 30 ($C(s,k)$ with $s=2,\dots,10$ and $k=20,25,30$.)

k	2	3	4	5	6	7	8	9	10
20	3.5	7.9	12.1	25.3	29.4	64.9	61.4	112.7	138.1
25	3.5	7.7	12.3	24.1	32.3	39.1	52.3	103.1	92.3
30	3.4	7.9	16.6	22.3	33.3	49.1	69.3	58.6	117.3

TABLE 2.7: Average ratio of the number of small squares to the number of biq squares for $(1,2),\dots(1,10)-$ square tilings of squares of sizes $20,25,30$ ($B_{number}(s,k)$ with $s=2,\dots,10$ and $k=20,25,30$.)

$$C(s_2) \geq 1 - \frac{C(s_1)-1}{C(s_1)\alpha - 1}.$$

□

2.3.2 1-dimensional case

Consider one-dimensional tilings of a block of size n by two type of tiles: of length 1 and of length $l+1$. Let $p_n(l)$ be the average ratio between the number of small blocks to the number of big blocks in tilings of stripes of length n. Let $a(n)$, $b(n)$ be the total number small/big blocks in tilings of a stripe of length n for a given l, then $p_n(l) = \frac{a(n)}{b(n)}$. We have that $a(n)+a(k) \leq a(n+k)$ and $b(n)+b(k) \leq b(n+k)$, so the limit of $\frac{a(n)}{b(n)}$ exists.

Let

$$p(l) = \lim_{n\to\infty} p_n(l). \tag{2.11}$$

Proposition 2.15. *The bigger the size of a block is, the less often blocks of this size occur. Formally: if $l_1 > l_2$, then $p(l_1) > p(l_2)$.*

Proof. Let $N(n,b_l)$ denote the number of tilings of the stripe of size n by a tiles of size 1 and b_{l-1} tiles of size l. Further we omit the index of b for simplification. We have that $N(n,b) = \binom{n-lb}{b}$. Let us see for which b the ratio $\frac{N(n,b+1)}{N(n,b)}$ is equal to 1. The solution

$b(n)$ will then give us the number of big tiles that maximises $N(n,b)$.

$$\frac{\binom{n-l(b+1)}{b+1}}{\binom{n-lb}{b}} = \frac{(n-bl-b)(n-bl-b-1)\dots(n-bl-b-l)}{(n-bl)(n-bl-1)\dots(n-bl-l+1)(b+1)}$$
$$= \frac{\prod_{i=0}^{l}(n-bl-b-i)}{(b+1)\prod_{i=0}^{l-1}(n-bl-i)} \tag{2.12}$$

Multiply the denominator and the nominator in (2.12) by $\frac{1}{n^{l+1}}$ and set them equal:

$$\prod_{i=0}^{l}\left(1-\frac{bl+b}{n}-\frac{i}{n}\right) = \frac{b+1}{n}\prod_{i=0}^{l-1}\left(1-\frac{bl}{n}-\frac{i}{n}\right). \tag{2.13}$$

If one solves (2.13) for b, then solution $b(n)$ is the number of big tiles that maximizes $N(n,b)$.

Define $\lambda_n(l) = \frac{n}{b_l(n)}$ and let n go to ∞. Denote $\lambda(l)$ as the limiting value of $\lambda_n(l)$. Then $(l+1)\times\frac{1}{\lambda(l)}$ is the average proportion of space occupied by tiles of size $l+1$ and $p(l)=\lambda(l)-l-1$. Then (2.13) becomes

$$\left(1-\frac{l+1}{\lambda(l)}\right)^{l+1} = \frac{1}{\lambda(l)}\left(1-\frac{l}{\lambda(l)}\right)^{l} \Leftrightarrow \tag{2.14}$$

$$(\lambda(l)-(l+1))^{l+1} = (\lambda(l)-l)^{l}, \tag{2.15}$$

$$p(l)^{l+1} = (p(l)+1)^{l}. \tag{2.16}$$

For $l>1$, on gets

$$p(l) \approx \frac{l}{\log l-(1+o(1))\log\log l}. \tag{2.17}$$

This proves that if $l_1>l_2$, then $p(l_1)>p(l_2)$. □

For example, one can explicitly solve (2.13) for $l=1$ (which corresponds to tilings by bricks of size 1 and 2) and get

$$\lim_{n\to\infty}\frac{b_1(n)}{n} = \frac{5-\sqrt{5}}{10} \approx 0.27,$$

$$p(1) = \lim_{n\to\infty} \frac{n - 2b_1(n)}{b_1(n)} = \frac{1+\sqrt{5}}{2} \approx 1.61.$$

For $l = 2$ (tiles of size 1 and 3), one gets

$$\lim_{n\to\infty} \frac{b_2(n)}{n} \approx 0.19,$$

$$p(2) = \lim_{n\to\infty} \frac{n - 3b_2(n)}{b_2(n)} \approx 2.14.$$

2.4 Topological entropy and approximate counting

We have already mentioned the interest behind the counting problem. Since there is not much hope of finding a closed formula for the number of configurations and finding any sort of formula has proved to be extremely challenging, the approximate counting starts to be of interest.

We consider the topological entropy of the square tilings. It measures exponential growth of the number of configurations of an $n \times n$ region. More formally, let $A_n(m, s)$ be the number of $(m, s)-$square tilings of an $n \times n$ region.

Definition 2.16. Topological entropy of an $(m, s)-$square tiling system of a square region of size n is

$$\xi_{top}(m, s) = \lim_{n\to\infty} \frac{\log A_n(m, s)}{n^2}.$$

An efficient approximation of the $(1, s)-$square tiling system to within $\frac{1}{n}$ can be done due to the fact that the square tiling system is a **block gluing** $\mathbb{Z}^2$ shift of finite type (SFT) (see, e.g., [65] for basic definitions of subshifts).

Definition 2.17 ([77])**.** A $\mathbb{Z}^d$ subshift X is called **block gluing with gap size** $g \in \mathbb{N}$ if for any pair of solid blocks $B_1, B_2 \subsetneq \mathbb{Z}^d$ with separation $\delta(B_1, B_2) > g$ and any pair of points $y, z \in X$ there exists a point $x \in X$ such that $x|_{B_1} = y|_{B_1}$ and $x|_{B_2} = z|_{B_2}$. X is **block gluing** if it is block gluing with gap size g for some $g \in \mathbb{N}$.

Proposition 2.18. *Any $(1, s)-$square tiling system is block gluing.*

Proof. For an $(1, s)-$square tiling system to be block gluing means to be able to glue two rectangular regions together along the side of size n for some gap g. It can be done by taking $g = 1$, and thus the gap region is a stripe of size $1 \times n$. □

Definition 2.19. Let $(t_n \in \mathbb{N})_{n\in\mathbb{N}}$ be a non-decreasing sequence. $r \in \mathbb{R}$ is **computable with rate** $(t_n \in \mathbb{N})_{n\in\mathbb{N}}$ if there exists a deterministic Turing machine which, for any $n \in \mathbb{N}$, calculates a rational approximation $r_n \in \mathbb{Q}$ of r such that $|r - r_n| \leq \frac{1}{n}$ in at most t_n steps.

As states the theorem for a block gluing SFT $\mathbb{Z}^2$ by Pavlov-Schraudner [77]:

Theorem 2.20 (Theorem 1.4 [77])**.** *For any block gluing $\mathbb{Z}^2$ SFT there exists a constant $C \in \mathbb{R}$ such that the topological entropy of the SFT is computable with rate $(2^{Cn^2})_{n\in\mathbb{N}}$.*

Let us apply Pavlov's theorem to square tilings. This gives directly the following result:

Corollary 2.21. *For a $(1, s)-$square tiling system there exists a constant $C \in \mathbb{R}$ such that $\xi_{top}(1, s)$ is computable with rate $(2^{Cn^2})_{n\in\mathbb{N}}$.*

Remark. In the next section we show that there are cases when the dynamics of square tilings in terms of Markov chains is fast mixing, then the entropy approximation can be done in polynomial number of steps.

2.5 Random generation

In this section we are going to address a similar range of questions in the topic of random generation as the ones considered for dimers. In subsection 2.5.1 we define a height function for the square tilings and local transformations (flips) following definitions and results by [18, 54, 55, 82].

We introduce a Markov chain and study its properties in subsection 2.5.3. Ideally, we want to be able to rigorously bound the time that it takes for the chain to reach stationarity (its mixing time). If it is polynomial (over the size of the tileable region that is denoted by R throughout this chapter), then for $(1, s)$ tilings one gets an efficient way to approximately count them (as was mentioned at the end of section 2.4). The analyze of the initial chain is challenging and we explain why it is so. We consider a weighted version (where weights λ are assigned to one type of squares in a tiling) in subsection 2.5.5. Adding weights helps to notice phase transitions in the system. Usually the weights λ are assigned in such a way that the case $\lambda = 1$ corresponds to the unweighted version of the chain. This case is generally hard to analyze but it becomes easier to analyze the dynamics for λ below and above some critical point. We draw polynomial bounds on the mixing time of the Markov chain in the case of $(1, s)-$tiling with a condition on λ.

For very thin and long regions (for example, rectangles of size $n \times \log n$) the initial chain turns out to be fast mixing in the $(1, s)$ case (see subsection 2.5.4). We conjecture that it is fast mixing for any m and s when considered for these long regions. We also present simulations and conjectures (subsection 2.5.6) that show that the mixing time of the unweighted version might not be polynomial but sub-exponential, and might be as difficult to analyze as the critical cases of Markov chains for the independent sets and perfect matchings. A relation to the independent sets is discussed in subsection 2.5.5.

At the end, in subsection 3.3.7, we conjecture an analog of the Aztec diamond for dominoes (hexagon for lozenges) and present simulations that show an "Arctic circle"–type phenomenon.

2.5.1 Height function

Height functions for tilings by rectangular tiles appear in numerous works (e.g., [18, 54, 55, 82]) that use Conway tiling groups [94]. Having a well defined height function permits to get a structure on the set of configurations and then use it to construct an algorithm that verifies the tileability of a given region.

Let a be a symbol corresponding to a horizontal step of length one to the right on the rectangular grid, $a^{-1}-$ to a step to the left, $b-$ to a vertical step up and $b^{-1}-$ to a step down. Then any path on the skeleton of any tiling of R is a word in the alphabet $\{a, a^{-1}, b, b^{-1}\}$, as well as the perimeter of R.

Consider a group G generated by a and b and restrictions that ensure that a path around each tile is elementary. A path around an $m \times m$ square is $a^m b^m a^{-m} b^{-m} =: [a^m, b^m]$. Similar for the other tile. Then the tiling group G is defined as follows: $G = \langle a, b \mid [a^m, b^m], [a^s, b^s] \rangle$, where $[a^m, b^m] = [a^s, b^s] = e$. Any path g on a skeleton of a tiling of R can be expressed using elements of G. For example, consider a tiling of a $2m \times 2m$ rectangle by $m \times m$ squares. Let l be the left lower corner of the region, r be its rights upper corner. Then $g(l, r) = a^m b^m a^m b^m$ is a path (one of many) from l to r. Another path is, e.g., $g'(l, r) = a^m a^m b^m b^m = a^{2m} b^{2m}$.

2.5.1.1 Quotient group H

In order to define a height function, we introduce a quotient group H of G. Let $H = G/\langle a^m, b^s \rangle$. Since $a^m = b^s = e$, then $a^m b^m = b^m a^m$ and $a^s b^s = b^s a^s$. So $H = \langle a, b \mid a^m, b^s \rangle = \mathbb{Z}_m \times \mathbb{Z}_s$ Let Γ_H be the Cayley graph of H. It is a tree made of cycles of size m with cycles of size s attached to every vertex. Example for $m = 2$, $s = 3$ and for $m = 3, s = 4$ are shown in Figures 2.12, 2.13.

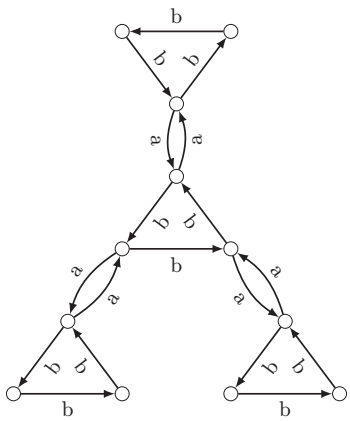

FIGURE 2.12: Cayley graph for $m = 2$, $s = 3$.

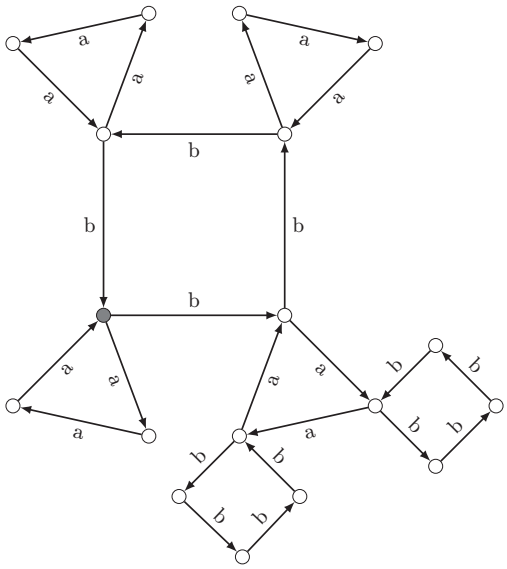

FIGURE 2.13: Cayley graph for $m = 3$, $s = 4$.

Every path g has a unique (canonical) expression using elements of $\mathbb{Z}_m \times \mathbb{Z}_s$:

$$g = k_1 l_1 k_2 l_2 \dots k_r l_r,$$

where $k_i \in \mathbb{Z}_m, l_i \in \mathbb{Z}_s$ for all $1 \leq i \leq r$ and some non-negative r. It can easily be constructed following the edges of the Cayley graph.

2.5.1.2 Weighted Cayley graph

Consider the Cayley graph Γ_H. Choose a vertex v_1, let it be the starting point and set its **weight** $w(v_1) = 0$. v_1 belongs to two cycles $C^m_{v_1}$ of size m and $C^s_{v_1}$ of size s. $C^m_{v_1} = \{v_1, v_2, \dots v_m\}$. Set

- $w(v_i) = w(v_1)$ for $i = 2, \dots, m-1$,

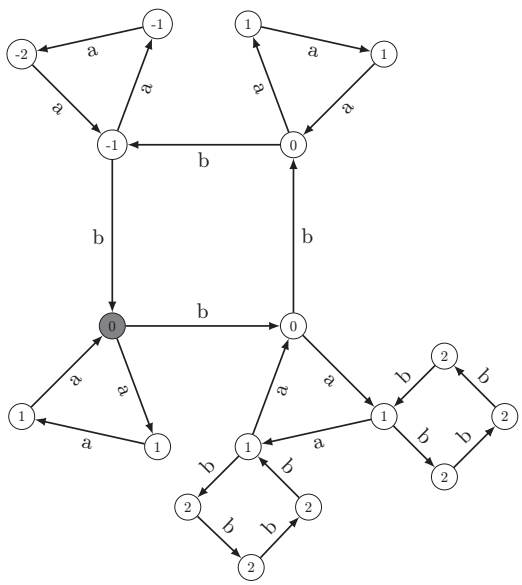

FIGURE 2.14: Weighted Cayley graph for $m = 3$, $s = 4$.

- $w(v_m) = w(v_1) - 1$.

Let $C^s_{v_1}, C^s_{v_2}, \ldots, C^s_{v_m}$ be cycles of size s that $v_1, \ldots v_m$ belong to accordingly. Set for all $i = 1, \ldots, m-1$

- $w(y) = w(v_i) + 1$ for all $y \in C^m_{v_i}$, $y \neq v_i$

Set for $i = m$

- $w(z_j) = w(v_m)$ for $z_j \in C^k_{v_m} \setminus \{v_m\}$ where $j = 1, \ldots m-2$,
- $w(z_{m-1}) = w(v_m) - 1$.

Continue assigning weights to the vertices of the Cayley graph in the way described above. The graph is infinite, a small part of the Cayley graph for $m = 3, s = 4$ is shown in Figure 2.14.

Another way to think about it is the following:

- In each cycle all vertices are of weight w, except for one that has weight $w - 1$, let us call it **descending**.
- Two cycles are connected via a vertex that is descending in one cycle and is not descending in the other.

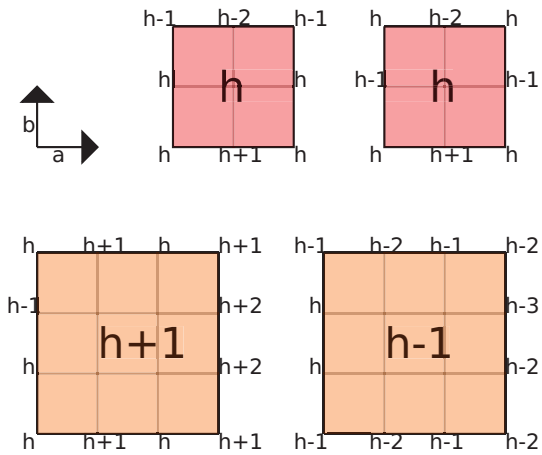

FIGURE 2.15: Heights for t_2 and t_3 tiles.

2.5.1.3 Height function for $(m, s)-$square tilings

For each $(m, s)-$ tiling T of a finite region R of $\mathbb{Z}^2$ denote the set of vertices of its skeleton by V_T. The height h is defined on V_T as follows. Choose an initial point x on the boundary of R (if R is a rectangle x can be chosen as the left lower corner of R for example). Without loss of generality, set $h(x) := w(v_1)$. To calculate the height of any given y, construct a path $g(x, y)$ connecting x and y in the following way: start the path in x and make moves that do not cross any tile. Every time a horizontal or vertical move is made, do the corresponding move in the weighted Cayley graph. The weight of the final node gives $h(y)$ and does not depend on the choice of the path. Let us point out that the height function on the boundary is completely defined by the region itself and does not depend on the tiling.

For $m, s \geq 2$ the height function h defined on vertices of $(m, s)-$ square tilings of a rectangular region R is flat on the boundary (it goes around the same two cycles on the Cayley graph) and $O(\text{size of } R)$ on the interior points. For the degenerate case when m or s equals one, h is simply a constant.

There are two tiles: t_s is the square tile of size s, t_m- of size m. Consider t_m. Since elements of $\mathbb{Z}_m$ encode the horizontal steps, the heights of vertices on one of the vertical sides are exactly the same as the heights on the other vertical side. let h be the maximum over heights of vertices on the vertical side, then the maximum over heights of vertices on each of the horizontal sides belongs to $\{h-1, h+1\}$. Define the height of the tile t_m $h(t_m) := h$. Change vertical sides to horizontal sides to get the similar property for t_s. Figure 2.15 shows how the heights are defined each node for t_2 and t_3 using the weighted Cayley graph where weights are used to write heights in each vertex.

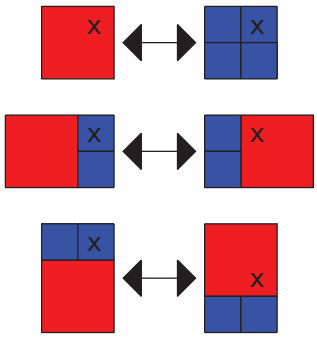

FIGURE 2.16: Flips for the $(1,2)-$ tiling.

Let $h(T)$ be the height of a tiling T. Define it as follows:

$$h(T) = \sum_{t \in T} a(t)h(t),$$

where $a(t)$ is the area of a tile t.

For the degenerate case thus defined $h(T)$ is not very interesting since $h(T) \equiv const * area(R)$ for every tiling T.

2.5.1.4 Flips

Let us define flips for $(m,s)-$square tilings. Call a **horizontal block** a rectangle of size $sl \times (m+s)$ where l is the least positive integer such that m divides sl. There are exactly two ways to tile this block. A **horizontal flip** is a flip in a horizontal block is a change from one tiling to another. In the same way, define a **vertical block** as a rectangle of size $(m+s) \times sl$ and let **vertical flip** be a flip in a vertical block. Call a **central block** a square of size $p \times p$, where where l is the least positive integer that divides m and s. For $(m,s)-$tilings, $m, s > 1$ there exists exactly two kinds of tilings of the block: by $(p/m)^2$ squares of size m and by $(p/s)^2$ squares of size s. A **central flip** is a flip in a central block. If $(m,s) = 1$, then $p = ms$. For $(1,s)-$tilings there are three tilings of the horizontal/vertical blocks. See Figures 2.16, 2.17 for an example of flips for the $(1,2)$ and $(2,3)$ tilings. The little black crosses in the figures mark the site that has to be chosen in order to perform a flip.

One can see from Figure 2.16 that there is no ambiguity about what kind of flip has to be performed for $(1,s)-$tilings in a horizontal/vertical block: if one wishes to perform a central flip, then the upper right corner of the big square has to be chosen, if one wishes to perform a horizontal flip, then one has to choose the site that will correspond to the upper right corner of the big square after it is moved.

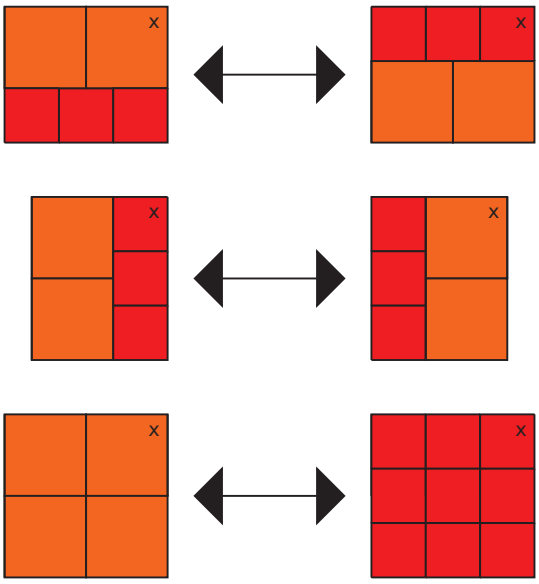

FIGURE 2.17: Flips for the $(2,3)-$ tiling.

It was shown in [82] that the tiling space is connected by flips. Moreover, there exists a minimal tiling to which one get by performing height non-increasing flips for every simply connected region. There is also an algorithm that runs in quadratic time over the size of the region that checks tileability by trying to construct a minimal tiling T_{min}. This minimal tiling might not be unique. There might be a subset of height equivalent minimal tilings $M = T_{min}$. In this case, one can consider a function on the subset of height equivalent tilings that (called **potential** in [82]). For $T_{min} \in M$ the potential adds up vertical coordinates of horizontal sides of the first type of tiles and horizontal coordinates of vertical sides of the second type of tiles. T_{min} that has the minimal potential is then the unique "global" minimal tiling on the set of tilings Ω_R.

Since there are exactly two tilings for every type of block, we say that each flip has two directions. One way to think about is the following: if a flip changes the height of the tiling, then the direction of the flip is "up" if it increases the height and "down" if it decreases the height. If the flip does not change the height, then it changes the potential. Let us then say that the direction of the flip is "up" if it increases the potential and "down" if it decreases the potential.

2.5.2 Examples

$(1,s)-$square tilings

Consider tilings by 1×1 and $s \times s$ squares of a finite region R of $\mathbb{Z}^2$ of area N_R. The height function defined above is not much of a use in this case, it is simply a constant.

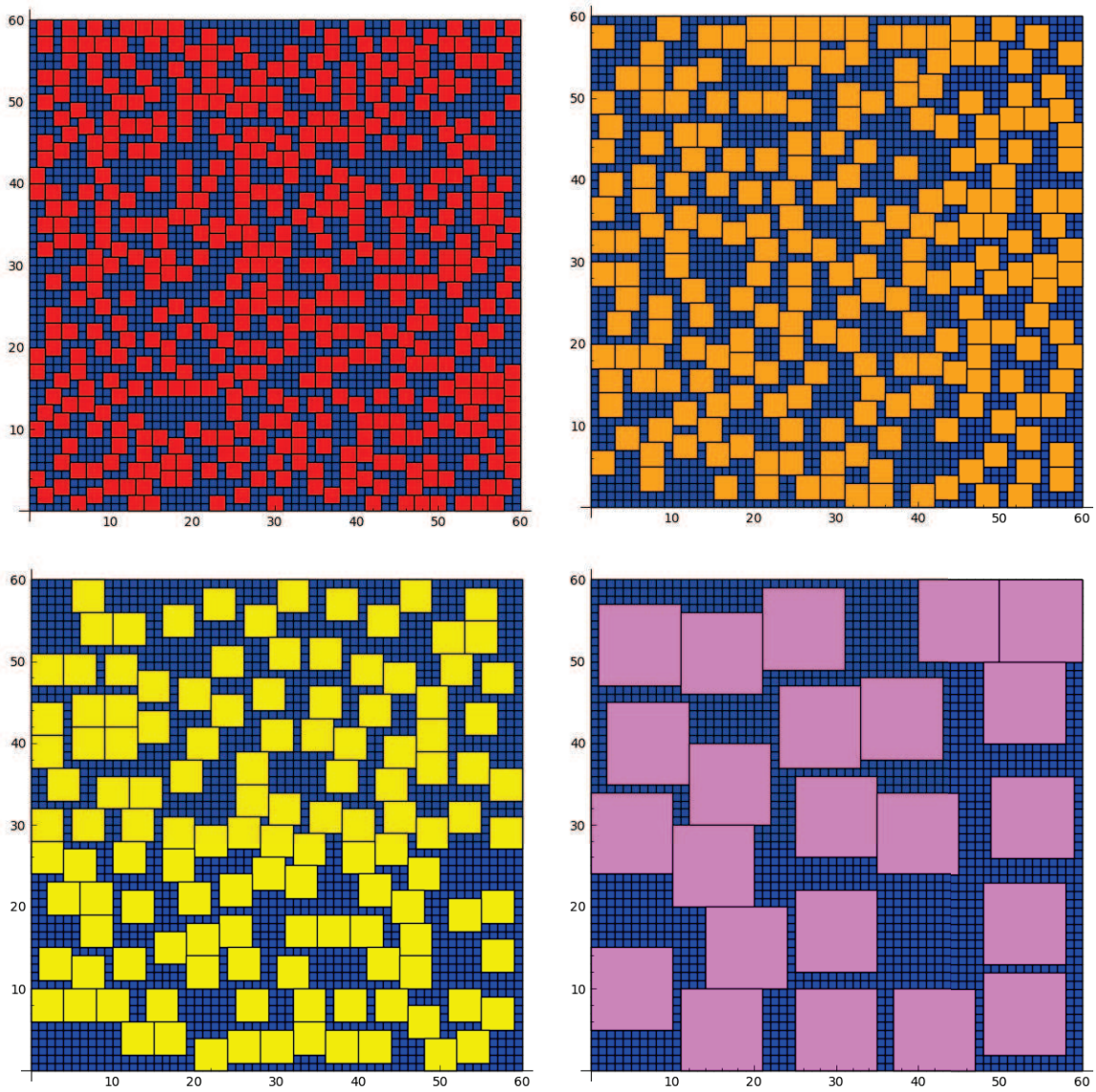

FIGURE 2.18: $(1,2),(1,3),(1,4),(1,10)$ tiling of a 60×60 square.

Consider central flips on the set of tilings of a region R. One can define a different order on the set of tilings. For example, for a given tiling T let its height $h(T)$ be equal to the minimal number of flips needed to get to from the tiling with only small squares T_0 to T, set $h(T_0) = 0$. The height function becomes simply a hamming distance function on the tiling graph $G_{1,s}(R)$. In this setting, T_0 is the unique minimal tiling. If the region R can be tiled by big squares only, then the maximal tiling is the tiling by big squares only. If speaking about rectangular regions, both of its sides need to have a multiple of s as a length.

If we consider central flips on the set of tilings of a region R, then the diameter of the tiling graph is $d(G_{1,s}) = O(N_R)$. Let α_s be the number of squares of size 1 in a tiling of

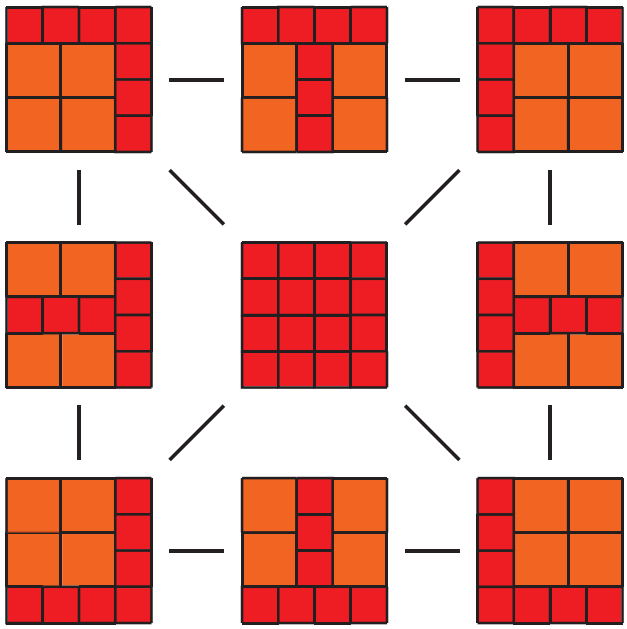

FIGURE 2.19: All $(2,3)-$square tilings of an 8×8 square.

R, then $\alpha_s = const(R) \mod s^2$. It does not depend on a tiling and holds for any finite region R of $\mathbb{Z}^2$.

Figure 2.18 shows examples of $(1,2), (1,3), (1,4), (1,10)$ tilings of a 60×60 region. Tiling by 1×1 and 2×2 squares is a random tiling obtained via coupling (see subsection 2.5.6 for details on simulations). The other three tilings are outputs after $100,000,000$ flips (starting from configurations with only 1×1 squares).

$(2,3)-$square tilings

$(2,3)-$square tilings (and (m,s) in general) are completely different from the $(1,s)$ case. It is no longer possible to glue two parts of a tiling together that easily. It depends a lot on the boundary of the region. The height function is linear over the size of the region. Figure 2.19 shows all possible $(2,3)-$square tilings of an 8 region where two configurations are connected if a flip can be made to go from one to the other. There are two ways of defining a height function and assigning weights. In one of the ways, the minimal tiling is in the upper left corner, the maximal is in the lower right corner.

2.5.3 Markov chains

Let R be a tiled simply connection region of area N, denote the set of all possible tilings by Ω_R. Let us define a Markov chain $\mathbf{MC}_{square}$ for square tilings. The idea of the chain is to pick a site of the region uniformly at random at each step and do a flip if possible.

First of all, for an $(m, s)-$tilings, $m, s > 1$, the site that has to be chosen to perform a flip is always in the upper right corner of the block. One can see from Figure 2.17 that once a site of a tiling is chosen, there is no ambiguity in what type of flips can be performed: in order to perform a flip, one has to consider the three blocks in which this site is in the upper right corner, and there is at most one type of flips that can be performed. If $m = 1$, one has to consider all blocks around the site in which it is the upper right 2×2 box, then there is at most one type of flips that can be made as well (which can be seen from Figure 2.16).

Second of all, in order to choose which of two configurations of the block we want (or in the case $m = 1$, in which direction we want to push the big square), let us recall that each flip has two directions, so before performing a flip we first choose one of two possible directions.

Let us now formally present the Markov chain:

MC$_{square}$:

Let $T_0 \in \Omega_R$ be an initial configuration. At each time t:

- choose an inner vertex of R with probability $\frac{1}{N}$,
- choose a direction with probability $\frac{1}{2}$,
- perform either a vertical/horizontal or central flip in the tiling T_t in the chosen direction if possible thus defining the tiling T_{t+1}, otherwise stay still.

Lemma 2.22. *MC$_{square}$ has uniform stationary distribution.*

Proof. The probability to reach every tiling $T \in \Omega_R$ is positive since the space state is connected, so MC$_{square}$ is irreducible. The probability to stay in the same state is positive, therefore it is aperiodic. This implies that the chain is ergodic and thus has unique stationary distribution. Moreover, the probability matrix of MC$_{square}$ is symmetric: indeed, $\mathbb{P}(T, S) = \mathbb{P}(S, T)$ for all pairs (T, S) from Ω_R (that are different by one flip). The symmetry of the probability matrix ensures the uniformity of the stationary distribution. □

2.5.4 Mixing time for $(1, s)-$square tilings

The question of proving fast mixing turns out to be difficult and there are only few tiling systems for which the dynamics was proven to be fast mixing, by fast mixing we

mean in polynomial number of steps in the size of the tiled region. We show that in some particular cases, the dynamics is rather easy to understand and is related to known structures such as independent sets.

Theorem 2.23. *MC_{square} is rapidly mixing for $(1, s)$ tilings of a rectangular region of size $n \times \log n$.*

Proof. We are going to use the **canonical paths** argument developed by Sinclair [89]. The idea of canonical paths is the following: fast mixing occurs when a tiling graph does not have a bottleneck. If there is a bottleneck, then it divides the set of configurations into two subsets. This slows down the mixing because it becomes hard to get from one subset to the other. Canonical paths allow to formalize absence (or presence) of a bottleneck. For each pair of configurations a canonical path (or a set of paths) is (are) defined that allow(s) to get from one to the other via flips. Consider an edge in the tiling graph (it connects two configurations different by one flip). If for any edge the number of paths that pass through this edge is relatively small (linear over the cardinality of the set of configurations), the graph does not have a bottleneck.

Let us present the construction for the $(1, 2)$ case. It is exactly the same for the general case. Place the rectangle on the grid with the lower left corner in $(0, 0)$. Define an order on the set of inner vertices of the $(n + 1) \times (\log n + 1)$ rectangle as follows:

$$\{(1, 1), (1, 2), \ldots, (1, \log n), (2, 1), (2, 2), \ldots, (n, \log n - 1), (n, \log n)\}$$

and denote them as $\{1, 2, \ldots \tilde{n}\}$, where $\tilde{n} = n \times \log n$.

Consider two tilings X and Y from Ω. In order to get from X to Y, it is sufficient to make a flip in every vertex of the region, in other words, the diameter of the tiling graph is not greater than the area of the region. With the use of canonical paths, we can structure the order in which we perform flips. Define a canonical path $p(X, Y)$ from X to Y as follows: start from X, follow the vertices in the defined order in windows of size 2×2 and in each window perform at most one flip in each vertex if it decreases the flip-distance (Hamming distance in the tiling graph) to Y. Such ordering of flips at each step either corrects a 2×2 box in X by performing a central flip or decreases distance by one by performing a vertical/horizontal flip in a vertical/horizontal block. After all vertices are met once, the path reaches Y. Due to the construction, for any pair of tilings such path exists and is unique. The ordered path is now a permutation σ of vertices $1 \ldots \tilde{n}$, where for each i : $\quad i - \log n - 1 \leq \sigma(i) \leq i + \log n + 1$.

Denote by π the chain's stationary distribution. Send $\pi(X)\pi(Y)$ units of **flow** through $p(X, Y)$ for any pair (X, Y) The flow through an edge is just the sum of all the flow

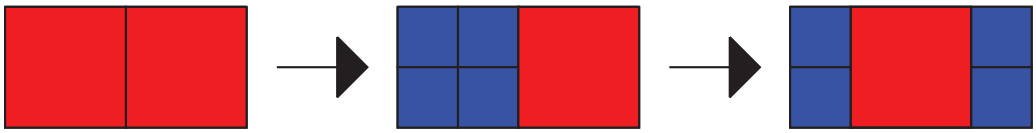

FIGURE 2.20: Canonical path for the $(1,2)-$ tiling in a 4×2 region.

that travels through the edge. The idea of the canonical paths method is to prove that any edge in the tiling graph has a small number of canonical paths going through it and therefore little flow. Consider an edge $e_i = (Z_{\sigma(i)}, Z_{\sigma(i+1)})$, where tilings $Z_{\sigma(i)}$ and $Z_{\sigma(i+1)}$ only differ in the $\sigma(i)-$th position.

Lemma 2.24. *There are no more than $|\Omega|n$ paths passing through each e_i.*

Proof. Consider the canonical path p going from X to Y:

$$p(X,Y) = \{X = Z_{\sigma(1)}, \dots, Z_{\sigma(i)}, Z_{\sigma(i+1)}, \dots Z_{\sigma(\tilde{n})} = Y\}.$$

Tiling X agrees with $Z_{\sigma(i)}$ at least on the vertices $\sigma(i+1), \dots, \sigma(\tilde{n})$. Y agrees with Z_i at least on the vertices $\sigma(1), \dots, \sigma(i)$. One could think of reconstructing X and Y using e_i and the first i vertices $1, \dots, \sigma(i)$ of X at the last $\tilde{n} - i$ vertices $\sigma(i), \dots \sigma(\tilde{n})$ of Y, but that would mean that it was possible to glue two pieces of a tiling with holes together. Consider instead a strip of width 4 around the windows that contain the vertex i between the two part of tilings which can be filled in at most $exp(\log n)$ ways.

We have therefore just constructed a map from the set of paths that pass through e_i $\{p_{e_i}\}$ to the state space Ω. This construction maps each path p_{e_i} to n tilings that are different only in the $4 \times \log n$ strip around the vertex i. Each path is mapped to a different "family" of $(1,s)-$ tilings, where each family corresponds to tilings of the strip around a given vertex. There are not more $|\Omega|$ possible families and therefore not more than $|\Omega|n$ paths through e_i.

□

Let us continue with the proof of the theorem. Applying Lemma 2.24, the flow along each e_i is at most $|\Omega|n$.

The **cost** of the flow f is

$$cost(f) := \max_{e_i} \frac{|\Omega| n \pi(X) \pi(Y)}{c(e_i)},$$

where X and Y are the endpoints of the paths that go through e, $c(e_i)$ is the **edge capacity**:

$$c(e_i) := \pi(Z_i)\mathbb{P}(Z_i, Z_{i+1}).$$

Since

$$\pi(X) = \pi(Y) = \frac{1}{|\Omega|}, \tag{2.18}$$

one gets the following bound on the cost:

$$cost(f) \leq 2n\tilde{n} = 2n^2 \log n. \tag{2.19}$$

There is the following relation between the cost function and the mixing time (see [89], Proposition 1 and Corollary 6'):

$$\tau_{mix}(\varepsilon) \leq 8cost^2(\ln \pi(X)^{-1} + \ln \varepsilon^{-1}). \tag{2.20}$$

Plugging (2.18) and (2.19) into (2.20), the following bound is obtained on the mixing time of MC_{square}:

$$\tau_{mix}(\varepsilon) \leq 32n^4 \log^2 n(c_2 n \log n + \ln \varepsilon^{-1}). \tag{2.21}$$

The same reasoning works for any $(1, s)-$ square tiling – the vertical strip has to be taken of length s. And τ_{mix} stays polynomial. □

Remark 1. Theorem 2.23 works not only for rectangular regions but for any regions, such that in any site the region can be divided in two parts via a strip of width $const(s)$ and height $\log n$.

Remark 2. Simulations via coupling (see subsection 2.5.6 for the description of the algorithm) suggest the $O(n^2)$ bound for the $(1, 2)$ case, which shows that the bound obtained in (2.21) is not optimal. Let us remind that the evident lower bound is of the size of the diameter of the tiling graph which is simply the area of the region.

2.5.5 Weighted Glauber dynamics

It is not clear how to prove fast mixing in the general case, so in this part let us consider a weighted version of the dynamics for the square tilings. It simply means that we favorise some configurations more than the others. In the case of $(1, s)-$tilings it seems

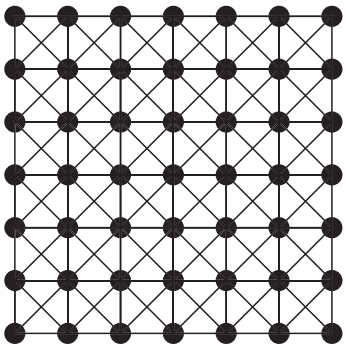

FIGURE 2.21: 8-regular graph on the square grid.

that the big squares significantly slow down the mixing time, so a way to go around is to put less probability weight on the biq squares. A Markov chain associated with the system that does local transformation, e.g. flips, (whose stationary distribution is the desired distribution) is generally referred to as the **Glauber dynamics**. It is popular in statistical physics and used to describe different behaviour of the system (Ising model, Hardcore model, independent sets, perfect matchings, etc.). Weights λ assigned to the particles usually correspond to the energy. Weights can help to detect presence of **phase transitions** in these systems – there exists a critical point λ_c such that the dynamics is fast mixing (in polynomial time over the size of the problem) below this critical point for all $\lambda < \lambda_c$ and is slow mixing (in exponential time) for all $\lambda > \lambda_c$. It is usually difficult to understand what is happening in the critical point. Let us just point out that that for a variety of studied models it has not been possible to analyze the behaviour of Markov chains in critical points and sometimes in their neighbourhoods.

We consider the $(1, s)-$case and prove fast mixing for certain $\lambda \leq \frac{1}{(2s-1)^2-2}$. In the case $s = 2$ the conditional bound on λ is better because of the relation to the independent sets.

2.5.5.1 $(1, 2)-$square tilings with weights

Consider $(1, 2)-$ square tilings as the King's problem on a toroidal region. If a king is placed in the site, none of 8 neighbouring sites can be occupied. Thus the King's problem can be seen as an independent set problem on the $8-$adjacency graph on the square grid graph $G = (V, E)$ of degree 8 (see Figure 2.21).

Consider the Glauber dynamics of this system. Let $I(G)$ be the set of all independent sets of G. An independent set of G is a subset of vertices such that no two of them are adjacent (for more information about independent sets see, e.g., [23, 25]). The

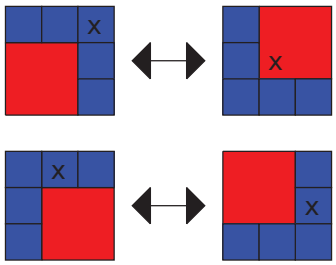

FIGURE 2.22: A diagonal drag flip for $(1,2)-$ tiling. x marks a site that has to be chosen to perform a flip (this site shows to where the big square is dragged).

probability of a configuration X is given by

$$\pi(X) := \frac{\lambda^{|X|}}{Z(\lambda)},$$

where λ is a positive parameter called a **weight** of a configuration and $Z(\lambda)-$ the **partition function** of the system:

$$Z(\lambda) := \sum_{X \in I(G)} \lambda^{|X|}.$$

Define the weighted version of $MC_{(1,2)-square}$ as follows. We add a **diagonal dragging** flip which is shown in Figure 2.22.

$MC_{(1,2)-square}$ with λ

Start from X_0. Let X_t be the configuration at time t. At time t:

- Choose a site of the region with probability $\frac{1}{N}$,
- If a central flip can be made, put four 1×1 squares with probability $\frac{1}{\lambda+1}$, put a 2×2 square with probability $\frac{\lambda}{\lambda+1}$. Else if a horizontal/vertical/diagonal (dragging) flip can be made, perform it with probability $\frac{\lambda}{4(\lambda+1)}$.
- Otherwise, do nothing and set $X_{t+1} = X_t$.

This version $MC_{(1,2)-square}$ corresponds exactly to the *delete/insert/drag* chain MC_{drag} for the independent sets [25], which is a slightly different version of the Luby-Vigoda chain [67].

MC_{drag} for independent sets

Start from X_0. Let X_t be the configuration at time t. At time t:

- Choose v from the set of vertices,
- If $v \in X_t$, then **delete** v : $X_{t+1} = X_t \setminus \{v\}$ with probability $\frac{1}{\lambda+1}$,
- If $v \notin X_t$, then **add** v : $X_{t+1} = X_t \cup \{v\}$ with probability $\frac{\lambda}{\lambda+1}$,
- If $v \notin X_t$ and v has a unique neighbour u in X_t, then **drag** v : $X_{t+1} = (X_t \cup \{v\}) \setminus \{u\}$ with probability $\frac{\lambda}{4(\lambda+1)}$,
- Otherwise, do nothing and set $X_{t+1} = X_t$.

The two main tasks are to approximately evaluate the partition function $Z(\lambda)$ and to approximately sample from $I(G)$ according to the stationary distribution. When $\Delta \geq 3$, approximate evaluation of $Z(\lambda)$ and approximate sampling from $I(G)$ can be done using a rapidly mixing chain (see, for example, [46]). Using the path coupling argument ([10] and Appendix A), Dyer and Greenhill proved fast mixing for MC_{drag} with sufficiently small λ. Rapid mixing for $MC_{(1,2)-square}$ then follows directly when $\lambda \leq \frac{1}{3}$.

Theorem 2.25 (Dyer-Greenhill [25])**.** *Let $G = (V, E)$ be a graph with maximal degree Δ and $|V| = n$. MC_{drag} is rapidly mixing for $\lambda \leq 2/(\Delta - 2)$.*

1. *When $\lambda < 2/(\Delta - 2)$,*

$$\tau_{mixing}(\varepsilon) \leq \frac{2(1+\lambda)}{2-(\Delta-2)\lambda} n \log(n\varepsilon^{-1}).$$

2. *When $\lambda = 2/(\Delta - 2)$,*

$$\tau_{mixing}(\varepsilon) \leq \lceil 2n^2(1+\lambda)(\log n + 1)\rceil \lceil \log \varepsilon^{-1} \rceil.$$

Using Dyer-Greenhill's theorem, we get the following bound on the mixing time for $MC_{(1,2)-square}$:

Theorem 2.26. *Consider $(1,2)-$square tilings of an $n \times n$ toroidal region. $MC_{(1,2)-square}$ is rapidly mixing for $\lambda \leq \frac{1}{3}$. The following bounds stand for $\tau_{mix}(\varepsilon)$ and some positive constant c_1, c_2.*

1. *When $\lambda < 1/3$,*

$$\tau_{mixing}(\varepsilon) \leq c_1 n^2 \log(n\varepsilon^{-1}).$$

2. *When $\lambda = 1/3$,*

$$\tau_{mix}(\varepsilon) \leq c_2 n^4 \lceil (\log n + 1) \log \varepsilon^{-1} \rceil.$$

2.5.5.2 $(1, s)-$square tilings with weights

Consider now the weighted dynamics for the $(1, s)$ case. Weights are put on the squares of size s (big squares). Consider the natural Markov chain with only central flips.

MC$_{(1,s)-square}$ with λ

Start from X_0. Let X_t be the configuration at time t. At time t:

- Choose a site of the region with probability $\frac{1}{N}$,
- If a central flip can be made – put s^2 1×1 squares with probability $\frac{1}{\lambda+1}$ or an $s \times s$ square with probability $\frac{\lambda}{\lambda+1}$, thus defining X_{t+1}
- Otherwise do nothing and set $X_{t+1} = X_t$.

For two tilings A and B of the region R by 1×1 and $s \times s$ flips let $\varphi(A, B)$ denote the minimal number of flips one has to perform to get from A to B. It follows directly from the definition of a flip that $\varphi(A, B) = \varphi(B, A)$ for any A, B.

It turns out that with λ sufficiently small, the coupling time for MC$_{(1,s)-square}$ is be polynomial. Namely, we get the following result.

Theorem 2.27. *Consider $(1, s)-$square tilings of an $n \times n$ region. MC$_{(1,s)-square}$ is rapidly mixing for $\lambda \leq \frac{1}{(2s-1)^2-2}$ and there exists a positive constant c such that*

$$\tau_{mix}(\varepsilon) \leq cn^4 \log(n\varepsilon^{-1}).$$

Proof. Consider a coupling $(A_t, B_t)_t$ and two configurations A_t and B_t at time t that are different by one flip, thus $\varphi(A_t, B_t) = 1$. We want to apply the coupling theorem dy Dyer and Greenhill and prove that $\mathbb{E}\Delta\varphi \leq 0$.

Consider $s = 2$. In the worst case scenario there are 8 bad sites that increase the distance between the two configurations and only one that decreases. A bad flip always implies putting a big square in the configuration with small squares. It is done with probability $\frac{\lambda}{N(\lambda+1)}$, where $N = n^2$ is the area of the region. The one good site decreases the distance for both direction of a flip: this is done with probability $\frac{\lambda}{N(\lambda+1)} + \frac{1}{N(\lambda+1)}$. So

$$\mathbb{E}(\Delta\varphi) \leq -\frac{1}{N} + \frac{8\lambda}{N(\lambda+1)} \tag{2.22}$$

This means that when the right part of (2.22) is not greater than 0, the chain is rapidly mixing. By solving the inequality one gets the condition on λ: $\lambda \leq \frac{1}{7}$.

In the general case the number of bad sites does not exceed $(2(s-1)+1)^2 - 1$, so

$$\mathbb{E}(\Delta\varphi) \leq -\frac{1}{N} + \frac{((2s-1)^2) - 1)\lambda}{N(\lambda+1)}. \tag{2.23}$$

$\mathbb{E}(\Delta\varphi) \leq 0$ whenever $\frac{((2s-1)^2-1)\lambda}{(\lambda+1)} \leq 1$. This is true when

$$\lambda \leq \frac{1}{(2s-1)^2) - 2}.$$

In order to apply the coupling theorem, we also need there to exist an $\alpha > 0$ such that $\mathbb{P}(\varphi(A_{t+1}, B_{t+1}) \neq \varphi(A_t, B_t)) \geq \alpha$. The inequality holds for $\alpha = \frac{1}{N}$.

Now we can freely apply the theorem and get the following bound on the mixing time of an $n \times n$ region:

$$\tau_{mix} \leq cD^2n^2,$$

where c is some positive constant, D is the diameter of the tiling graph. Since it is $O(n^2)$, one gets the desired bound on τ_{mix}.

□

Remark. Simulations with weight parameter λ from Theorem 2.27 suggest coupling in $O(n^2)$ steps.

2.5.6 Simulations

Let us describe the algorithm used for simulations. We use cython to run the simulations and sage.graphics for the pictures

The algorithm **1** describes a basic coupling approach for getting a sample of a $(m, s)-$square tiling of a given regions. The main problem though is which configurations choose as initial configurations of the coupling. In the sense of the height functions, there is a maximal and minimal tiling and the distance between them equals the diameter of the tiling graph, so one can use them as the two starting points. The number of steps after which they meet gives an estimate on the coupling time and an idea on the general look of the tiling. Table 2.8 shows estimates on the average coupling time over 100 trials for

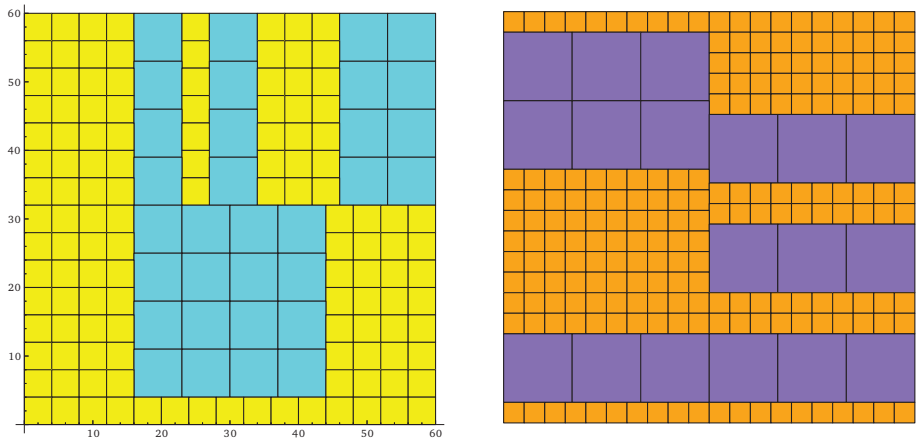

FIGURE 2.23: $(4,7)$ and $(3,10)-$square tilings of a 60×60 square via coupling.

$(1,s)-$tilings of a $n \times n$ region for $s = 2, \ldots, 10$ and $n = 10, \ldots, 30$. See Figure 2.23 for examples of $(4,7)$ and $(3,10)-$tilings obtained by coupling.

Algorithm 1 Coupling for $(m,s)-$square tilings of an $n \times n$ square

INPUT: $n-$area of the square, m and $s-$ sizes of squares

```
A ← Max tiling
B ← Min tiling
time t ← 0
while A ≠ B do
    direction ← random choice (0,1)
    position ← random choice of vertex
    A ← flip(A, direction, position)
    B ← flip(B, direction, position)
    t ← t + 1
end while
```

Conjecture. *Let R be an $n \times n$ square region of area n, tiled by squares of size m and s, $m < s$. Then $MC_{(1,2)-square}$ is rapidly mixing and*

$$\tau_{mix}^{(1,2)} = O(n^{3.5})$$

Moreover, for $s > 2$ $MC_{(m,s)-square}$ is not rapidly mixing its mixing time $\tau_{mix}^{(m,s)}$ is sub-exponential and has the following bound:

$$\tau_{mix}^{(m,s)} = O(exp([n/s] + 1)).$$

$n \setminus s$	2	3	4	5	6	7	8	9	10
10	1.2×10^3	2.7×10^3	1.01×10^3	4×10^2	2×10^2	1.6×10^2	3×10^1	1×10^1	1
11	2.9×11^3	5.02×11^3	$2. \times 11^3$	1.01×11^3	1.3×11^3	1.3×11^2	2×10^2	$1.3 \times 10^{1.5}$	0.8×10^1
12	$2.2 \times 12^{3.5}$	0.8×12^4	3.1×12^3	2.1×12^3	1.3×12^3	0.9×12^3	3×10^2	2.5×10^2	$2.1 \times 10^{1.5}$
13	$2.3 \times 13^{3.5}$	1.1×13^4	$2.6 \times 13^{3.5}$	3.1×13^3	1.1×13^3	1.24×13^3	$1.169 \times 13^{2.5}$	$1.7 \times 13^{2.5}$	1.01×13^2
14	$1.5 \times 14^{3.5}$	1.2×14^4	0.9×14^4	5.1×14^3	2.3×14^3	1.01×14^2	0.9×14^3	0.83×14^3	1.73×14^2
15	$1.1 \times 15^{3.5}$	1.9×15^4	1.2×15^4	8.2×15^3	4.1×15^3	1.1×15^3	1.7×15^3	1.3×15^3	$1.9 \times 15^{2.5}$
16	$1.1 \times 16^{3.5}$	2.5×16^4	1.9×16^4	0.9×16^4	$1.01 \times 16^{3.5}$	0.8×16^3	1.95×16^3	2.2×16^3	$3.9 \times 16^{2.5}$
17	$2.5 \times 17^{3.5}$	3.01×17^4	2.9×17^4	0.9×17^4	$2.5 \times 17^{3.5}$	3.1×17^3	2.6×17^3	2.82×17^3	0.95×17^3
18	$2.4 \times 18^{3.5}$	4.9×18^4	$1.01 \times 18^{4.5}$	1.3×18^4	0.8×18^4	1.9×18^3	3.5×18^3	3.3×18^3	1.7×18^3
19	$2.9 \times 19^{3.5}$	5.1×19^4	$1.4 \times 19^{4.5}$	2.01×19^4	0.85×19^4	2.5×19^3	$1.1 \times 19^{3.5}$	$0.89 \times 19^{3.5}$	2.2×19^3
20	$2.9 \times 20^{3.5}$	1.01×20^4	$2.9 \times 20^{4.5}$	3.7×20^4	1.1×20^4	1.2×20^3	$1.18 \times 20^{3.5}$	$1.32 \times 20^{3.5}$	2.5×20^3
21	$2.9 \times 21^{3.5}$	1.4×21^4	$1.8 \times 21^{4.5}$	$2.6 \times 21^{4.5}$	1.6×21^4	$1.01 \times 21^{3.5}$	$1.8 \times 21^{3.5}$	$1.7 \times 21^{3.5}$	$0.7 \times 21^{3.5}$
22	$2.3 \times 22^{3.5}$	1.8×22^4	$1.2 \times 22^{4.5}$	$2.2 \times 22^{4.5}$	3.01×22^4	3.2×22^4	$3.2 \times 22^{3.5}$	$2.1 \times 22^{3.5}$	$0.8 \times 22^{3.5}$
23	$2.9 \times 23^{3.5}$	1.2×23^5	0.7×23^5	0.8×23^5	$1.01 \times 23^{4.5}$	1.02×23^3	1.23×23^4	$2.9 \times 23^{3.5}$	$1.7 \times 23^{3.5}$
24	$2.2 \times 24^{3.5}$	0.7×24^5	$1.8 \times 24^{4.5}$	0.6×24^5	0.7×24^5	$1.7 \times 24^{3.5}$	1.4×24^4	1.05×24^4	$2.2 \times 24^{3.5}$
25	$2.9 \times 25^{3.5}$	1.4×25^4	$2.5 \times 25^{4.5}$	0.7×25^5	1.01×25^5	2.5×25^4	1.8×25^4	1.79×25^4	0.77×25^4
26	$2.3 \times 26^{3.5}$	2.7×26^5	1.1×26^5	0.7×26^5	1.1×26^5	1.8×26^4	2.6×26^4	3.2×26^4	1.28×26^4
27	$2.9 \times 27^{3.5}$	2.5×27^5	1.5×27^5	3.1×27^5	$1.2 \times 27^{4.5}$	$1.5 \times 27^{4.5}$	$1.02 \times 27^{4.5}$	2.8×27^4	1.2×27^4
28	$2.9 \times 28^{3.5}$	3.1×28^5	2.7×28^5	2.3×28^5	0.7×28^5	$1.8 \times 28^{4.5}$	$1.3 \times 28^{4.5}$	$1.2 \times 28^{4.5}$	1.7×28^4

TABLE 2.8: Average coupling time for $(1, s)$ tilings of $n \times n$ squares over 10 to 100 trials.

2.5.7 Limit shape

Let us consider a region such that the height function is not a constant on its boundary but rather grows linearly. One can think of an "Aztec diamond-type" region. It seems that in this case, tiling by squares of sizes m and s, if both $m, s > 1$ have a typical limiting look, similar to the Arctic circle for dimer tilings. An example is shown in Figure 2.24. We consider an hexagonal region with with a staircase border in the bottom and top part and flat in the middle. Each stair is of horizontal size s and vertical m, such that the height function becomes linear over the length of the side.

When $m = 1$, $s > 1$ there is no long-range property, since 1×1 squares can fit everywhere, so the shape of the region does not force any specific placement of tiles, see Figure 2.25.

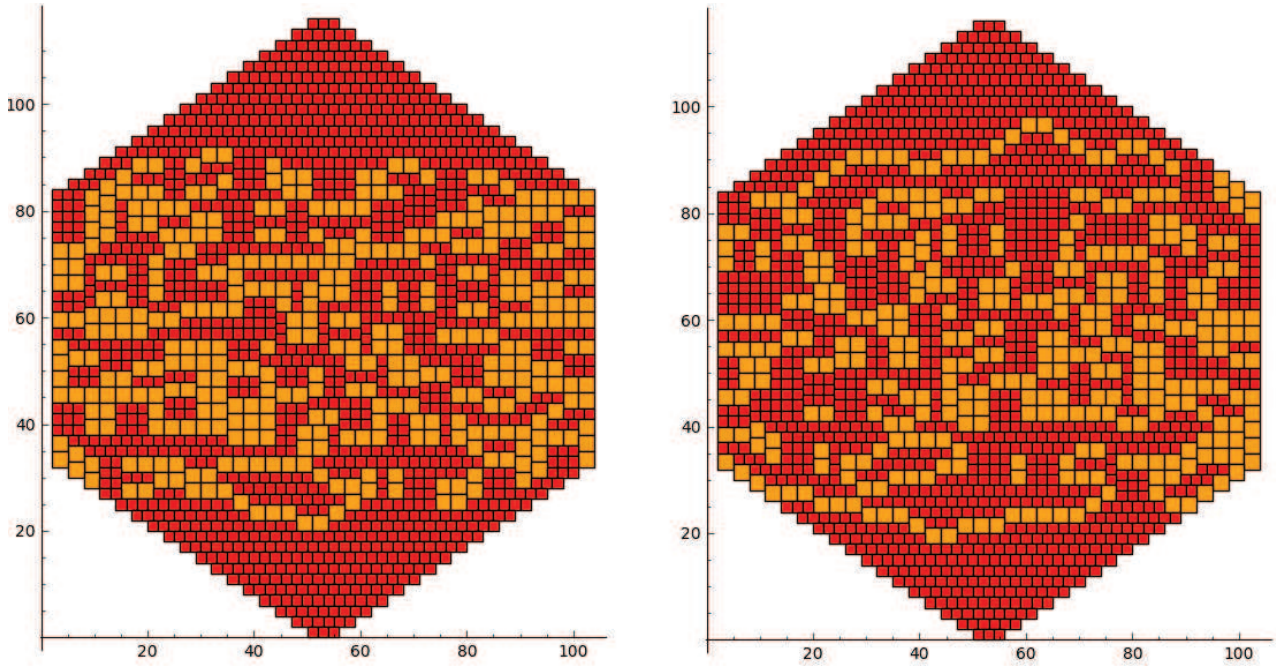

FIGURE 2.24: $(2,3)-$tiling of the 50-diamond after 10 million flips (left) and 50 million flips (right).

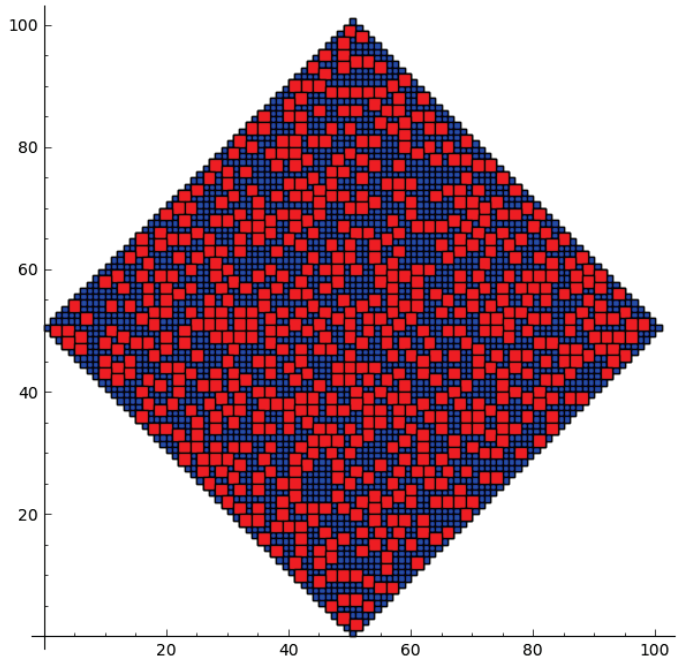

FIGURE 2.25: $(1,2)-$tiling of a diamond-shaped region.

Chapter 3

Kagome tilings

3.1 Introduction

In this chapter we consider tilings on the **Kagome** lattice (also known as the **Butterfly/tri-hexagonal** lattice) which is a partition of the plane into hexagons and triangles of the same side length such that each edge is shared by one triangle and one hexagon (see Figure 3.1). The dimer model on the tri-hexagonal lattice was studied along with other dimer models on regular lattices [98].

A prototile on the Kagome lattice is an hexagon with two adjacent triangles. there are three type of prototiles that we call **trapeze, fish** and **lozenge** (Figure 3.4). A closed region R of the Kagome lattice is tileable if there is at least one partition of the cells of this region into prototiles. An example of a Kagome tiling is shown in Figure 3.2.

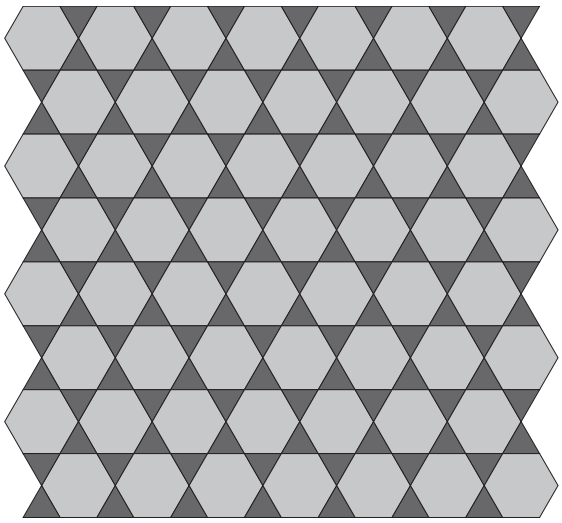

FIGURE 3.1: Kagome lattice

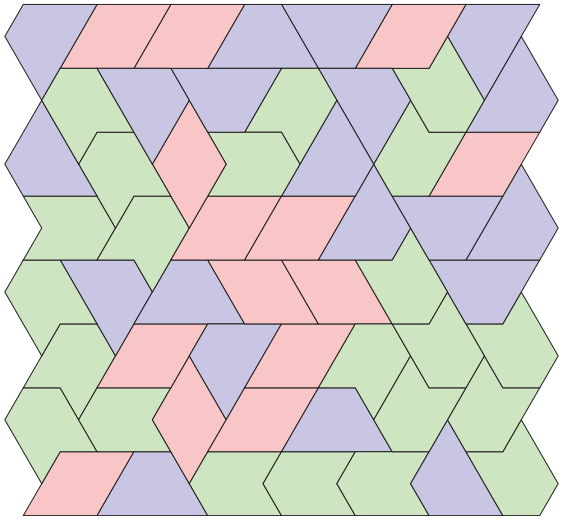

FIGURE 3.2: General Kagome tiling

We follow almost the same logic in this chapter as in Chapters 1 and 2 and try to answer questions related to counting and sampling. In section 3.2 we show a relation between Kagome tilings on a toroidal region and perfect matchings of a 6-regular bipartite graph, which allows to do approximate counting as soon as the graph satisfies a certain condition. In section 3.3 we focus on random generation, we introduce the height function as it was done in [6]; the connectivity of the configurational space under flips (that was proven in [6] as well) allows us to introduce a Markov chain and analyze its convergence time to stationarity. Even though simulations using Coupling from the Past (see Appendix B) suggest mixing time to be $O(N^{2.5})$, where N is the number of tiles, a rigorous proof seems to be rather challenging. In subsection 3.3.4 we explain why it is so. In subsection 3.3.5 we introduce a weighted version of the chain, as we did in Chapter 2 for square tilings. This allows to bound mixing time of the chain for a specific condition on weights λ (Theorem 3.4). In subsection 3.3.6 we also consider restrained Kagome tilings with only two prototiles (lozenge and trapeze). It turns out that for a lozenge-shaped region the restrained set of flips give the connectivity of the configurational space. We prove that the Markov chain on the restrained Kagome tilings is fast mixing for lozenge-shaped regions (Theorem 3.7)

3.2 Approximate counting

3.2.1 Kagome tilings as perfect matching

Consider a region R of the Kagome lattice with free boundary conditions. Let us associate a graph G_R to it in the following way. Put a node in the center of each triangle,

two nodes in the center of each hexagon. Connect each node of a triangle adjacent to an hexagon to two nodes of the hexagon as shown in Figure 3.3. G_R is a 6−regular graph.

There are exactly two matchings associated with each tile as shown on Figure 3.4. Each tiling with m tiles corresponds to 2^m possible perfect matchings, since for every tile there are two possible matchings and the choice of the matching is independent for each tile. Let Ω_R be the set of Kagome tilings of R, M_{G_R} be the set of perfect matchings of G_R. Then

$$|\Omega_R| = 2^m |M_{G_R}|.$$

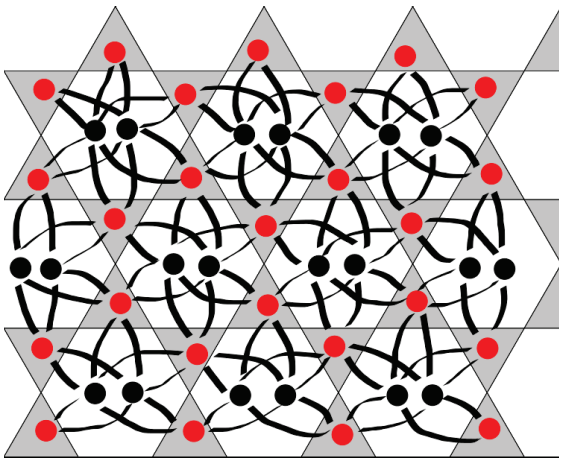

FIGURE 3.3: Bipartite graph associated to the Kagome lattice drawn on it.

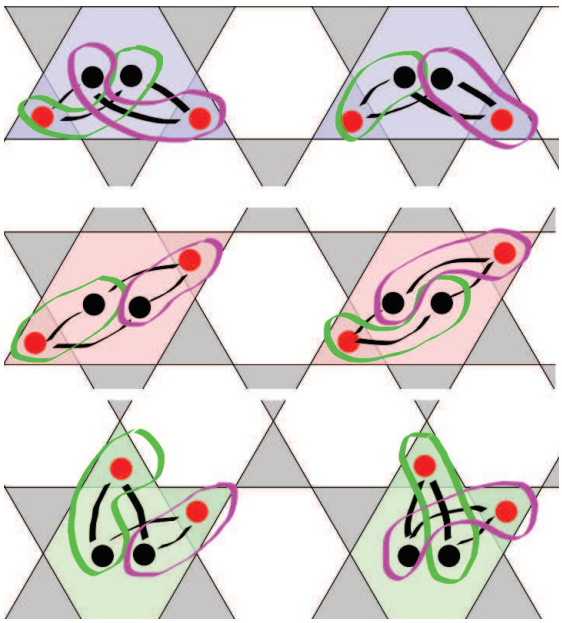

FIGURE 3.4: Two possible matchings for Kagome tiles (from top to bottom: trapeze, lozenge, fish).

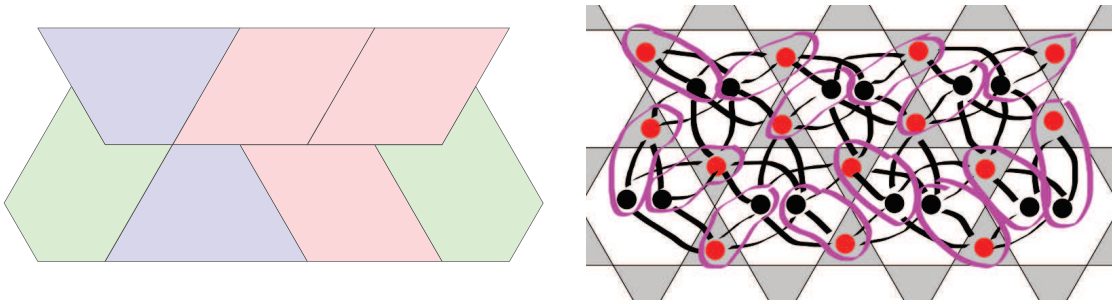

FIGURE 3.5: A tiling and one of possible perfect matchings associated with it.

One can approximately count $|\Omega_R|$ by approximately counting the number of perfect matchings $|M_{G_R}|$. We give a brief overview on approximate counting of perfect matchings in bipartite graphs in the following subsection.

3.2.2 Perfect matchings in bipartite graphs

The general question of counting (perfect) matchings belongs to the class of the $\#P-$complete problems [97]. Technically, in [97] Valiant proved the $\#P-$completeness for the computation of the permanent of a matrix with non-negative entries. When A has only entries 0 and 1, it can be seen as the adjacency matrix of a bipartite graph G_A, so its permanent then equals the number of perfect matchings in G_A. We have already seen that for planar graphs exact counting of perfect matchings can be done efficiently [52]. Note that exact counting of all matchings in planar graphs is however $\#P-$complete [43].

Consider a graph $G = (V, E)$ with $|V| = 2m$. Let M_G be the set of perfect matchings, N_G be the set of **near-perfect matchings** – matchings with exactly $m - 1$ edges. Let $\alpha(G) := \frac{N_G}{M_G}$ be the ratio of near-perfect matchings to perfect matchings. Note that $\alpha(G) \geq m$, since erasing one of m edges in each perfect matching gives a near-perfect matching.

Jerrum and Sinclair [45] proved that there exists an **fpras** (fully polynomial random approximation scheme, see Appendix B) for the number of perfect matchings in any $2m-$vertex graph G that satisfies $\alpha(G) \leq q(m)$ for a fixed polynomial q. The running time of fpras depends a lot on $\alpha(G)$, so for it to run fast, one would want $\alpha(G)$ to be not much bigger than m. In general, for bipartite graphs, the computation of $\alpha(G)$ is $\#P-$complete as proved in [5]. Kenyon, Randall and Sinclair proved in [59] that $\alpha(G)$ is essentially small for families of graphs with sufficiently strong symmetry properties More precisely, they proved that for any bipartite vertex-transitive graph G with $2m$ vertices $\alpha(G) \leq m^2$. This gives an efficient way to approximately count perfect matchings.

Jerrum, Sinclair and Vigoda later [47] showed how to approximately count perfect matchings in any bipartite graph using a Markov chain (that they refer to as the switch chain) with weights on the set near-perfect and perfect matchings in $O(m^{10}(\log m^3))$ steps. The counting problem of matchings is extensively discussed in [70].

3.2.3 Back to Kagome

Let R be a tileable region of the Kagome lattice with periodic boundaries by n tiles, G_R be the 6−regular bipartite graph associated with it with $2m$ vertices, set $m := 2n$. Then the following statement holds for G_R. Let M be the set of perfect matchings, for vertices $u, v \in N$ let $N(u, v)$ the set of near-perfect matchings with holes in u and v. Finally, let N be the union of $N(u, v)$ over all pairs of u, v: $N = \cup N(u, v)$.

Proposition 3.1.

$$\alpha(G_R) \leq m^2.$$

Proof. Let us prove that $|M| \geq |N(u, v)|$ for any $u, v \in V$. If it is true, then

$$|M|m^2 \geq |N_{u,v}|m^2 \geq |N|$$

The proof basically repeats the proof of Theorem 2, [59]. Let v' be a neighbour of v, such that the graph distance $d(u, v') > d(u, v)$, let u' be a neighbour of u. $N(u, v)$ is the set of near-perfect matching with holes in u and v, $N(u', v')$ – in u' and v' respectively. Let $e_u = (u, u')$ and $e_v = (v, v')$. Colour edges of each matching in $N(u, v)$ in black, edges of $N(u', v')$ and e_u, e_v in white. Consider $C = N(u, v) \cup N(u', v') \cup \{e_v, e_u\}$. Each $N \in C$ has a full cycle (u, v, u', v') of edges coloured in white and black such that each vertex belongs to one white and one black edge, except for u', v' that belong to two white edges.

Let us prove that $|N(u, v)||N(u', v')| \leq |M|^2$ by constructing an injective map $\psi : N(u, v) \times N(u', v') \longrightarrow M \times M$. The map works as follows: take $(N_1 \in N(u, v), N_2 \in (N(u', v'))$ such that their overlapping cover all edges in the cycle u, v, v', u' except for the edges e_v and e_u. Together with e_u, e_v they cover the cycle completely. Start from u' towards v', color the first edge in white, then black, etc, the edge e_v will be then coloured in black since the path from u' to v' has an odd length. Continuing like this from v to u, one gets e_u covered in black as well. This gives two different perfect matchings of the cycle. This defines the map ψ. ψ is injective simply because if one considers two perfect matchings from the image and does all the operations backwards, one gets two unique near-perfect matchings $N_1 \in N(u, v)$ and $N_2 \in N(u', v')$.

□

3.3 Random generation

Once again, after dimers and square tilings, we are interested in questions related to random generation. First we define a height function for the Kagome tilings and local transformations as it was done in [6, 94] see subsection 3.3.1. We introduce a Markov chain and study its properties in subsection 3.3.3. We are as always interested in finding bounds on the mixing time. Note that good bounds on the mixing time would allow not only to sample from the stationary distribution but also to do approximate counting which was discussed in section 3.2. For the Kagome tilings simulations show fast mixing, namely in $O(n^5)$ for a tileable region of size $2n$ (n^2 tiles)(see Table 3.1). It proves to be challenging to analyze the initial chain and we shortly explain why it is so in subsection 3.3.4. As we did for square tilings in the previous chapter, we consider a weighted version of the chain in subsection 3.3.5 and are able to draw polynomial bounds with a condition on λ. In subsection 3.3.6 we consider tilings by only two specific types of tiles instead of three and consider a version of the initial chain. Fast mixing then can be proved when the state space (the set of all possible configuration) is connected. We show that in case of lozenge-type regions, the state space is connected and the chain is rapidly mixing. In the end of the section (in subsection 3.3.7), we show some simulations done for a lozenge-type region that suggest a specific limit shape similar to the Arctic circle for dimer tilings.

3.3.1 Height function

Following Thurston [94], the orientation on the edges of the lattice is clockwise on triangles and anti-clockwise on hexagons. We assign $+1$ to an edge if it belongs to a tile and -2 otherwise (*flow* in [6]). If we follow the edges around a tile and sum the flows – with $+$ if we follow the orientation of arrows, with $-$ if not, – then the flow gives 0 around each tile. Let us now introduce the notion of **height**. We denote the set of lattice vertices in R by V_R and $N := |V_R|$. Fix a vertex $v \in V_R$ and let its height $h(v) := 0$. For any $w \in V_R$ its height $h(w)$ is defined by the flow from v to w. See Figure 3.6. A vertex $v \in V_R$ is called a **local minimum (maximum)** if $h(v)$ is less(greater) than $h(w)$ for all w that share an edge of the lattice with v. For a tiling T_R (or simply T) let its height be

$$h(T) := \sum_{v \in V_R} h(v).$$

n	t
10	2.91×5^5
12	2.53×6^5
14	1.92×7^5
16	1.91×8^5
18	1.76×9^5
20	1.43×10^5
22	1.44×11^5
24	1.35×12^5
26	1.31×13^5
28	1.15×14^5
30	1.07×15^5
32	1.13×16^5
34	1.05×17^5
36	1.05×18^5
38	1.04×19^5
40	1.02×20^5

TABLE 3.1: Coupling time for the Kagome tilings of a square region of size $2n$, $n = 5, \ldots 20$.

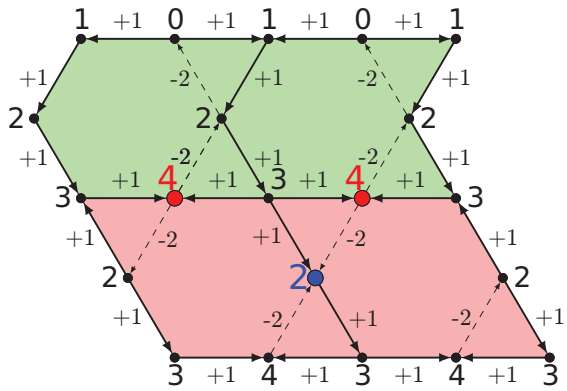

FIGURE 3.6: Flow, orientation and heights of a Kagome tiling.

3.3.2 Flips

Let us have two tiles a and b. Let a consist of an hexagon h_a and two triangles t_a and t_1, b of h_b, t_b and t_2, where t_a is adjacent to h_b, t_b is adjacent to h_a. Then the transformation of a and b into a', b' such that $t_a \in b'$, $t_b \in a'$ is called a **(simple) flip** (as defined in [6]). Flips can be performed only around local minima and maxima. A flip turns a local minimum into a local maximum and the other way around. We say that it has two directions: **minimal** if it decreases the height in the vertex and and **maximal** if it increases. All possible flips (up to rotations) are shown in Figure 3.7. An example of flips is shown in Figure 3.8.

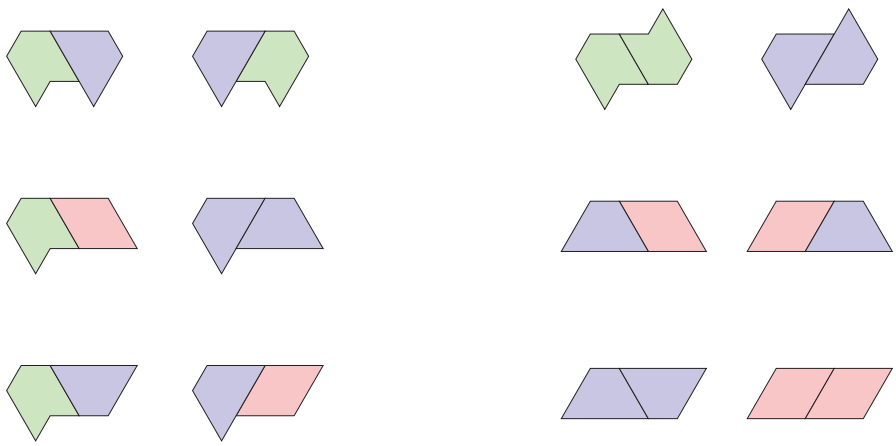

FIGURE 3.7: All possible flips for the general Kagome tiling.

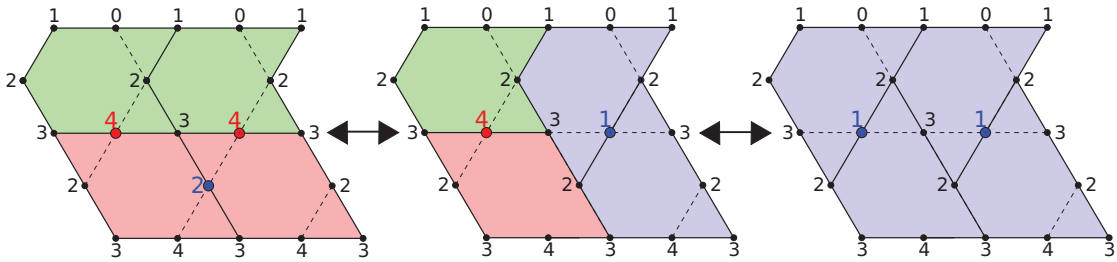

FIGURE 3.8: Example of flips performed in a Kagome tiling.

Proposition 3.2. *Let $n_h(v)$ be the number of different heights h in a vertex $v \in V_R$. Then for all $v \in V_R$*

$$n_h(v) \leq \sqrt{N}.$$

Proof. In order to prove this, let us introduce the following subsets of vertices of our system: we denote $V_{\partial R}$ by V_0 which contains all vertices that belong to the border ∂R. We consider that R has perimeter $\sqrt{N}$. The border does not change, so $n_h(v) = 1$ for $v \in V_0$. Now take vertices in $V \setminus V_0$ that have vertices from V_0 as their neighbours, and denote this subset by V_1. Now $n_h(v) \leq 2$ for $v \in V_1$. Take the subset of vertices from $V \setminus (V_0 \cup V_1)$ that have neighbouring vertices from V_1 and denote it by V_2, etc. R has a finite number of vertices, so in the end we will obtain the subset V_l. Since R has area N, $l \leq \sqrt{N}$. For all $(v, w) \in (V_i, V_{i+1})$, $i = 0, ..., l-1$: $|n_h(v) - n_h(w)| \leq 1$, then $n_h(v) \leq \sqrt{N}$.

□

Let G_R be the graph in which each vertex corresponds to a tiling of R, and two vertices are connected by an edge i.f.f. the corresponding tilings differ by one flip. If the graph has one connected component, there is a unique minimal tiling T_{min} (having the minimal height) and a unique maximal tiling T_{max} (having the maximal height) [6]. It follows from Proposition 3.2 that the diameter of the tiling graph is $D \leq N^{\frac{3}{2}}$, where $N = |V_R|$.

3.3.3 Markov chain

Let us define a Markov chain $\mathbf{MC}_{kagome}$ on a set of Kagome tilings of the region R with N inner vertices.

$\mathbf{MC}_{kagome}$:

Let T_0 be an initial configuration. At each time t:

- choose an inner vertex of R with probability $\frac{1}{N}$,
- choose a direction of a flip with probability $\frac{1}{2}$,
- perform a flip in the chosen direction in the tiling T_t if possible thus defining the tiling T_{t+1}, otherwise stay still.

Lemma 3.3. *MC_{kagome} has uniform stationary distribution.*

Proof. MC_{kagome} is *irreducible* since any tiling from Ω can be obtained from any other tiling via a finite number of consecutive flips, it is *aperiodic* since the self-loop probability is greater than zero. Therefore, the chain is *ergodic* and it converges to its unique stationary distribution. Moreover, the transition probabilities are symmetric, so the stationary distribution is uniform. □

3.3.4 Coupling of general Markov chain

We construct a coupling for MC_{kagome}. Let A_0 and B_0 be two starting configurations. At time t we pick a vertex from V and a direction (same for both tilings). We make a flip in A_t and B_t if possible; this defines A_{t+1} and B_{t+1}. This randomizing process respects the order: that is if $h(B_t) < h(A_t)$, then $h(B_{t+1}) \leq h(A_{t+1})$. Simulations using Coupling from the Past method (see Appendix B) for the MC_{kagome} suggest that its mixing time $\tau_{mix} = O(N^2)$.

We wish to apply the path coupling method (see Appendix A) to bound τ_{mix}. For tilings $A, B \in \Omega$ define the **distance** function φ on $\Omega \times \Omega \to Z$ as

$$\varphi^*(A, B) := h(A) - h(B),\ \varphi := \frac{\varphi^*}{3}.$$

Let U be a subset of $\Omega \times \Omega$ that consists of pairs of configurations that differ by one flip. We wish to prove that for all A_t, $B_t \in U$

$$\mathbb{E}[\Delta\varphi(A_t, B_t)] \leq 0,$$

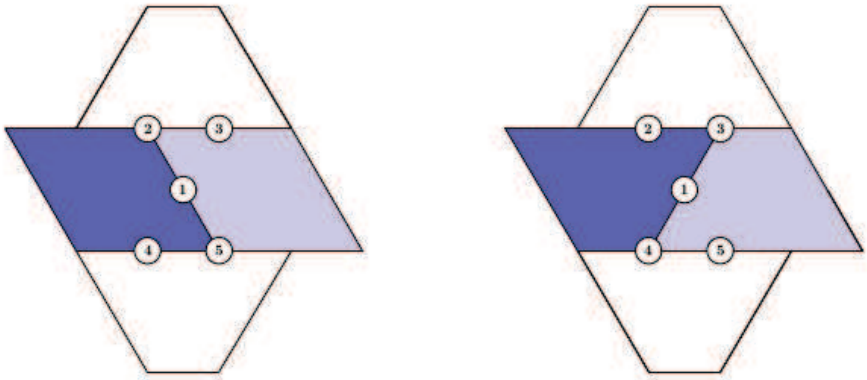

FIGURE 3.9: Bad configuration for path coupling.

where $\Delta\varphi(A_t, B_t) = \varphi(A_{t+1}, B_{t+1}) - \varphi(A_t, B_t)$.

Unfortunately, with such definition of the distance, the inequality does not stand. The following example illustrates that. Consider two tilings A and B that differ by one flip around the vertice 1 (see Figure 3.9).

$$\mathbb{E}[\Delta\varphi] = -\frac{1}{N} + \frac{1}{2N} + \frac{1}{2N} + \frac{1}{2N} + \frac{1}{2N} > 0. \tag{3.1}$$

This means that either a better metric should be found (for example, what Wilson did with the lozenge tilings [99]) or the Markov chain should be modified in such a way that would make it possible to analyse it (one can think of the tower of flips for domino and lozenge tilings by Luby, Randall, Sinclair [66]). In the next two subsections we consider a weighted version of the Markov chain such that the coupling method works, and we consider as well restrained Kagome tilings and prove that the corresponding Markov chain is fast mixing for specific regions.

3.3.5 Weighted Glauber dynamics

We have already seen in Section 3 how putting weights on certain configurations can make the analysis of Markov chains significantly easier. Let us do use the same approach for the Kagome tilings. We change the probabilities of flips in the following way. Consider Figure 3.7. Notice that there are 3 flips that make the *fish* tile appear/disappear, call them the **fish changer** flips, call others **fish free** flips. Modify the general MC in the following way by putting weights on the fish changer flips.

MC$_{kagome}$ with weight λ:

Let T_0 be an initial configuration. At each time t:

- choose an inner vertex of R with probability $\frac{1}{N}$,
- 1. If a fish changer flip can be made in the tiling T_t, make it with probability $\frac{1}{1+\lambda}$ if it makes a fish tile disappear and with probability $\frac{\lambda}{1+\lambda}$ if it makes a fish tile appear. This defines the tiling T_{t+1};
  2. If a normal flip can be made in T_t, perform it in either direction with probability $\frac{1}{2}$, thus defining T_{t+1};
  3. Otherwise stay still.

Remark. If $\lambda = 1$, one gets the initial MC$_{kagome}$.

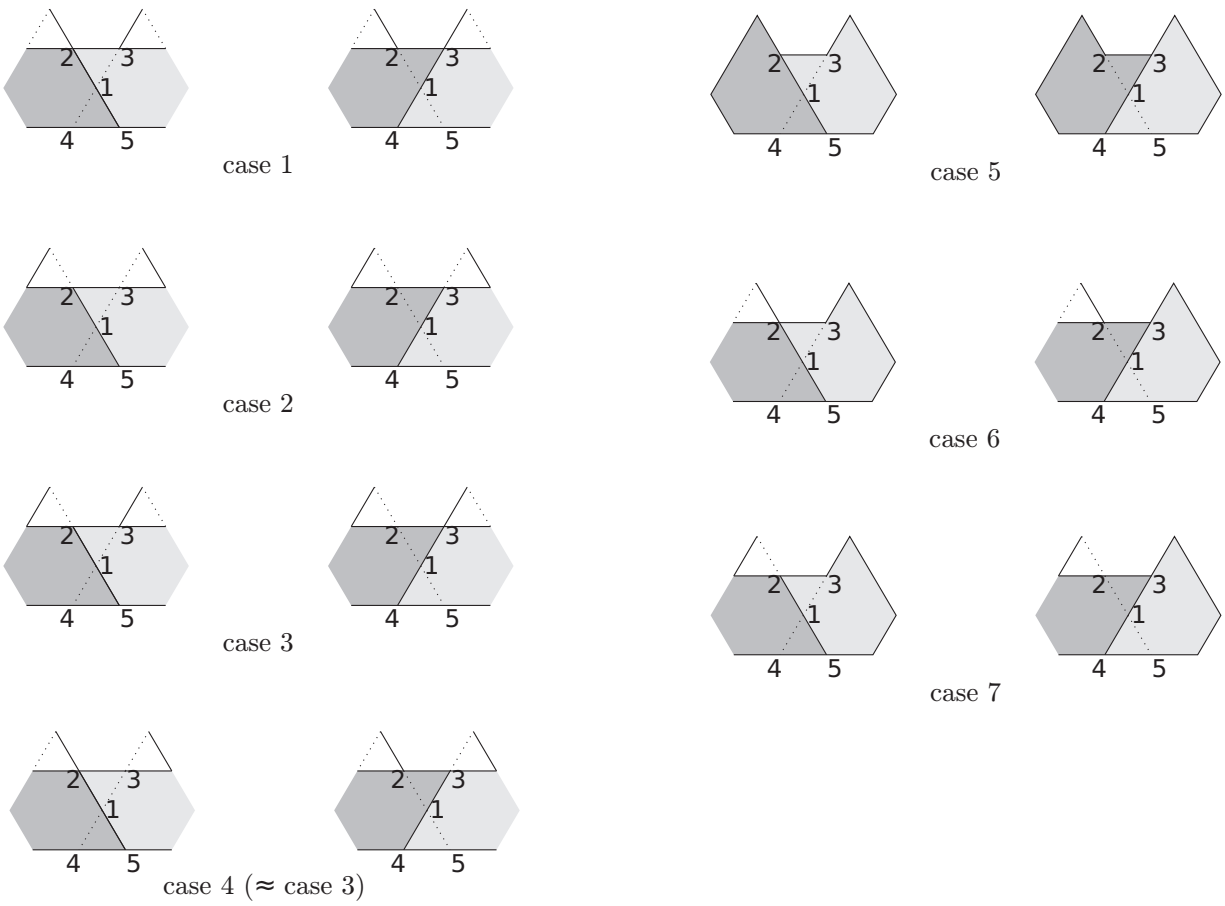

FIGURE 3.10: All possible cases for a pair of configurations at distance 1 are shown schematically.

Theorem 3.4. *Consider Kagome tilings of a region R of area N. MC$_{kagome}$ with weight λ is rapidly mixing for $\lambda \leq \frac{1}{3}$. The following bounds stand for τ_{mix} and some positive constant c:*

$$\tau_{mix} \leq cN^4 \tag{3.2}$$

Proof. Consider a pair of tilings (A, B) different by one flip. One can think of Figure 3.9 from the previous subsection. All possible configurations are shown in Figure 3.10. One observation is that if the vertex 1 is chosen, flips around it will decrease the distance by 1 in any of the case: in terms of expectation one will have either $-\frac{1}{2} - \frac{1}{2}$ if a fish free flip is made or $-\frac{1}{\lambda+1} - \frac{\lambda}{\lambda+1}$ if a fish changer flip is made.

Another observation is that there are at most 4 **bad** vertices – vertices that increase the distance between A and B. The only layout possible to have 2 bad vertices on each side (vertices 2,3 and 4,5) is to have a trapeze tile sharing its longest side with the two tiles in question as is shown in Figure 3.9. In this case a flip around any of the vertices 2,3,4 and 5 that can increase the distance is the fish changer flip that makes a fish tile appear. So it is performed with probability $\frac{\lambda}{\lambda+1}$. The equation (3.1) that failed mixing now turns into:

$$\mathbb{E}[\Delta\varphi] = -\frac{1}{N} + \frac{1}{N}\frac{\lambda}{\lambda+1} + \frac{1}{N}\frac{\lambda}{\lambda+1} + \frac{1}{N}\frac{\lambda}{\lambda+1} + \frac{1}{N}\frac{\lambda}{\lambda+1}. \tag{3.3}$$

We want (3.3) to be ≤ 0, so one gets :

$$\frac{4\lambda}{\lambda+1} \leq 1,$$

which yeilds the bound on λ

$$\lambda \leq \frac{1}{3}.$$

We are pretty much done with the proof, since the case with one bad vertex on each side gives, in terms of expectation, at most: $-\frac{1}{N} + \frac{1}{2N} + \frac{1}{2N} = 0$. And the mixed situation with one bad vertex on one side and two bad vertices on the other side gives, in terms of expectation, at most: $-\frac{1}{N} + \frac{1}{2N} + \frac{1}{N}\frac{\lambda}{\lambda+1} + \frac{1}{N}\frac{\lambda}{\lambda+1}$. This is surely less than 0 when $\lambda \leq 1/3$.

Since $\mathbb{P}(\varphi(A_{t+1}, B_{t+1}) \neq \varphi(A_t, B_t)|A_t, B_t) \geq \frac{1}{N}$, the coupling theorem gives

$$\tau_{mix} \leq c\frac{D^2}{1/N},$$

where D is the diameter of the tiling graph. So one gets (3.2). □

3.3.6 Restrained Kagome tilings

Consider Kagome tilings where the set of prototiles is restrained to a trapeze and a lozenge and flips using them – Figure 3.11. Note that partial order is still respected: for

a pair of tilings T_1, T_2 of the same region, if $h(T_1) \geq h(T_2)$, then after performing a flip in vertex v in both tilings (if possible) $h(T_1') \geq h(T_2')$.

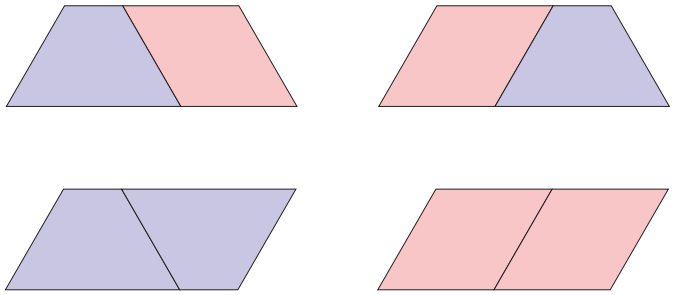

FIGURE 3.11: Possible flips for the restrained Kagome tiling.

Define a Markov chain MC^R_{kagome} in the same way as MC_{kagome} where a flip is performed only if it is one of the two allowed flips shown in Figure 3.11.

Throughout this part we consider that the region R is a lozenge of size $2n$. Denote the set of all such possible restrained Kagome tilings of R by Ω_{loz}. Now, consider a tiling $T_o \in \Omega_{loz}$. Start performing flips that decrease the height of the chosen vertex. Note that the order of flips does not matter since a flip in a vertex cannot block a flip in the same direction in neighbouring vertices (vertices that belong to the same triangle as the chosen vertex). Perform only decreasing flips while it is possible. In the tiling T_{min} obtained in the end all vertices are local minima apart from some that are local maxima (but not global) in which a restrained flip cannot be performed. This implies that $h_{T_{min}}(v) \leq h_{T_o}(v)$ and $h_{T_{min}}(v) \leq h_{\overline{T_o}}(v)$ for all $v \in R$ where $\overline{T_o} \in \{$Tilings achievable from T_o by flips $\}$. Call such a tiling T_{min} **minimal**. Let Ω_{min} be the set of all minimal tilings for R, $\Omega_{min} \subset \Omega_{loz}$. Note that for each tiling T_o there exists a unique tiling $T_{min} \in \Omega_{min}$ obtained by the sequence of height decreasing flips as described above.

Let us prove that $|\Omega_{min}| = 1$, so that there is exactly one minimal tiling T_{min} which depends only on the height of the boundary of R. Below is the algorithm that constructs a unique minimal tiling using only the height of the boundary. The construction produces a unique minimal tiling, so that for any $T \in \Omega_{loz} : h_{T_{min}}(v) \leq h_T(v)$ for all $v \in R$.

Algorithm for constructing the minimal tiling

Fix a lozenge region R with the height function h defined on ∂R.

1. Choose a starting vertex on the boundary. Start following the boundary anti-clockwise. Let v_1 be the first global maximum of the boundary (In the lozenge they are all on the horizontal sides). Place the tile that includes the hexagon on

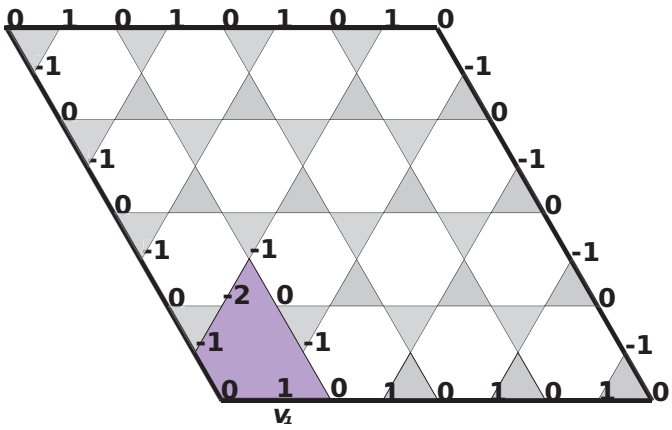

FIGURE 3.12: First step of constructing a minimal tiling starting from v_1 on the boundary.

the left from v_1 (See Figure 3.12) such that the global maximum is still achieved only on the boundary, if possible. If not, break.

2. Delete the tile and adjust the heights;

3. Go to the first step if the region is not empty after step 3. Otherwise, quit.

Remark 1. At step 1 there is at most one way of putting a tile because of the geometric properties of the tiles and the way in which the heights are defined for the points of the lattice. The placed tile that has the minimal height among all other possible tiles that include the hexagon on the left from it.

Remark 2. The algorithm will either break because a tile cannot be put, or will end when all the region is tiled and the tiling it produces is the minimal tiling (e.g., see Figures 3.13 and 3.14).

Remark 3. The way global maxima are chosen at step 1 can be changed in any other way. The tiling one obtains in the end does not depend on the order in which global maxima are chosen at each step due to the geometric properties of the tiles and the region. Moreover, it has the following properties: its global maxima occur only on the boundary of R, it does not have any flippable local maxima and its height is minimal among all possible tilings of R.

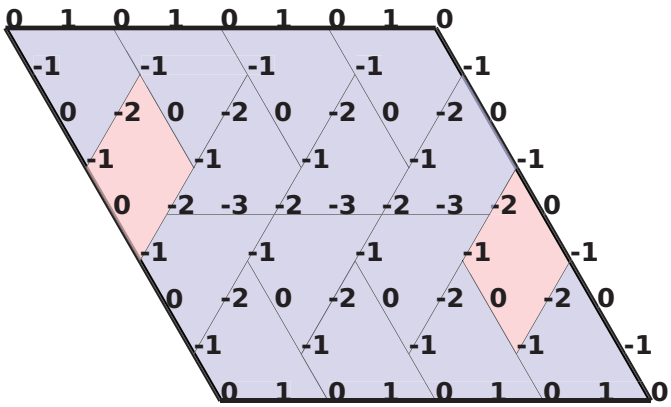

FIGURE 3.13: Minimal restrained Kagome tiling.

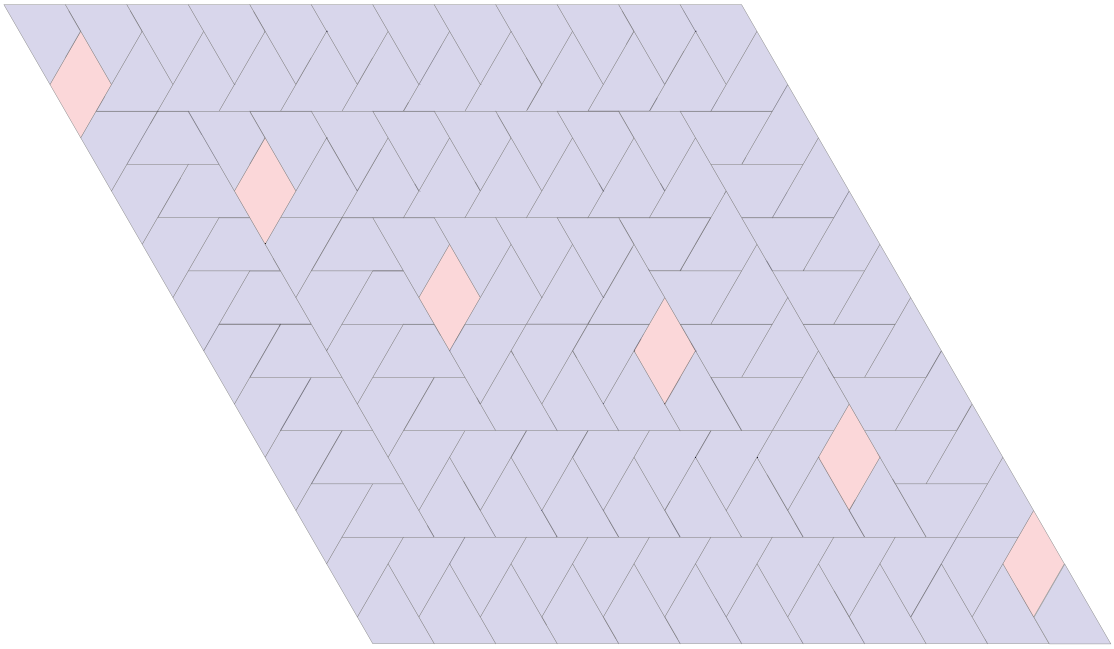

FIGURE 3.14: Minimal restrained Kagome tiling of a lozenge of size 24.

Proposition 3.5. *Let R be a tileable lozenge region and Ω_{loz} be the set of all possible Kagome tilings of R. Then Ω_{loz} is connected.*

Proof. We are going to prove that any two $T_1, T_2 \in \Omega_{loz}$ are connected by a sequence of flips. Let us put together a couple of facts that we have observed:

1. For each $T \in \Omega_{loz}$, due to the definition of the height function, the height h on ∂R depends only on R and not on the tiling.

2. For each $T \in \Omega_{loz}$ there exists a unique tiling $T_{min} \in \Omega_{min}$ achievable from T by height-decreasing flips.

3. h on ∂R defines a minimal tiling in a unique way.

It follows from $1-3$ that for any pair of $T_1, T_2 \in \Omega_{loz}$, both of T_1, T_2 are connected to the same minimal tiling T_{min}, and so they are flip-connected.

□

Remark. The choice of a lozenge region is due to the fact that there are simply connected regions for which Ω_{loz} is not connected. An example is shown in Figure...

Lemma 3.6. *MC^R_{kagome} has uniform stationary distribution.*

Proof. MC^R_{kagome} is *irreducible* since any tiling from Ω can be obtained from any other tiling via a finite number of consecutive flips, it is *aperiodic* since the self-loop probability is greater than zero. Therefore, the chain is *ergodic* and it converges to its unique stationary distribution. Moreover, the transition probabilities are symmetric, so the stationary distribution is uniform. □

Theorem 3.7. *Let R be a tileable region by the restrained family of Kagome prototiles of area N. Then MC^R_{kagome} is rapidly mixing. More precisely, there exists $c > 0$ such that the mixing time τ^R_{mix} of MC^R_{kagome} satisfies*

$$\tau^R_{mix}(\varepsilon) \leq cN^4 \lceil \ln \varepsilon^{-1} \rceil.$$

Proof. Consider restrained Kagome tilings and MC^R_{kagome}. We prove that MC^R_{kagome} is rapidly mixing by the use of the path coupling theorem that did not seem to work for the general case (see subsection 3.3.4).

Remember that the reason why it did not work for the general case, was that there were 4 vertices that increased $\Delta\varphi$ as show (3.1) and Figure 3.9). Now in this case of Figure 3.9, for a pair of tilings A and B that differ by one flip around the vertice 1, no flips can be made around vertices $2, 3, 4, 5$ which gives

$$\mathbb{E}[\Delta\varphi] = -\frac{1}{N}. \tag{3.4}$$

The other cases give at most one bad vertex on each side as shown in Figure 3.15. This gives $+\frac{1}{2N}$ in vertices 2 and 4, thus in total

$$\mathbb{E}[\Delta\varphi] \leq -\frac{1}{N} + \frac{1}{2N} + \frac{1}{2N} = 0. \tag{3.5}$$

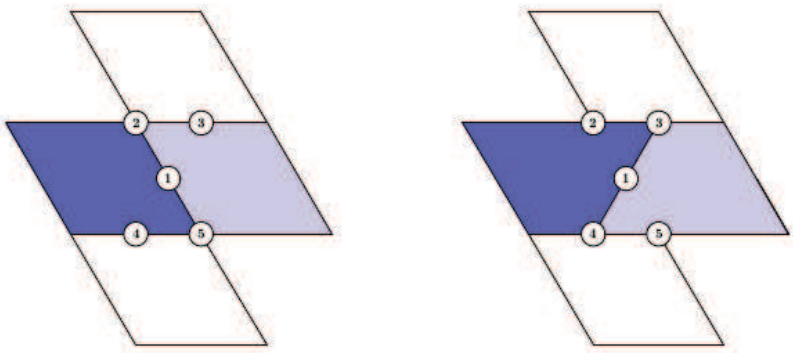

FIGURE 3.15: The worst case for the path coupling for MC^R_{kagome}.

If $A_t \neq B_t$ for all t, then

$$\mathbb{P}[\varphi(A_{t+1}, B_{t+1}) \neq \varphi(A_t, B_t)|A_t, B_t] \geq \frac{1}{N},$$

because in any case choosing the vertex 1 will decrease the distance. The coupling lemma gives the following bound:

$$\tau^R_{mix}(\varepsilon) \leq \frac{D^2}{\frac{1}{N}} \lceil \ln \varepsilon^{-1} \rceil. \tag{3.6}$$

Since for the restrained tilings, the diameter D is also $O(N^{\frac{3}{2}})$ (in the same way as for the general Kagome tiling using Proposition 3.2) and plugging it in (3.6) yields the desired bound:

$$\tau^R_{mix}(\varepsilon) \leq cN^4 \lceil \ln \varepsilon^{-1} \rceil,$$

where c is a positive constant.

□

3.3.7 Limit shape

Consider a lozenge region with a non flat boundary in such a way that the height function of the boundary is $\Omega(n)$, where n is the size of the region (See Figure 3.16). Then there appears to be a phenomenon similar to the Arctic circle in case of dominoes and lozenges that was discussed in Chapter 1. Figures 3.17, 3.18 [1] show simulations done using Coupling from the Past algorithm. It would be interesting to characterize the limiting shape rigorously.

[1] Figures 3.17, 3.18: courtesy of Th. Fernique

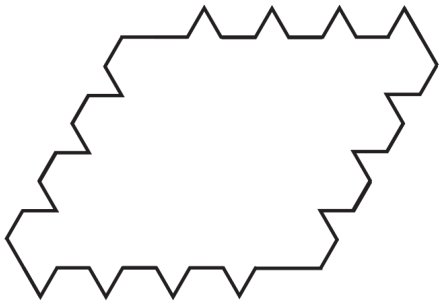

FIGURE 3.16: Kagome lozenge-shaped region with a non flat boundary.

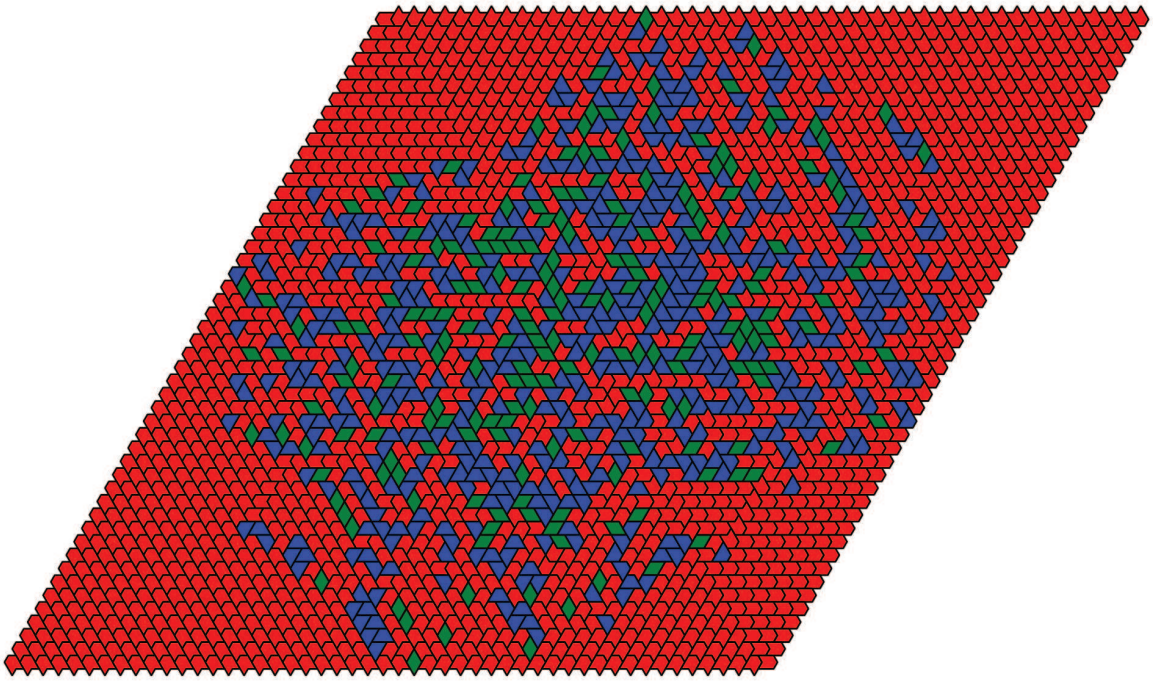

FIGURE 3.17: Kagome tiling of a lozenge of size 50.

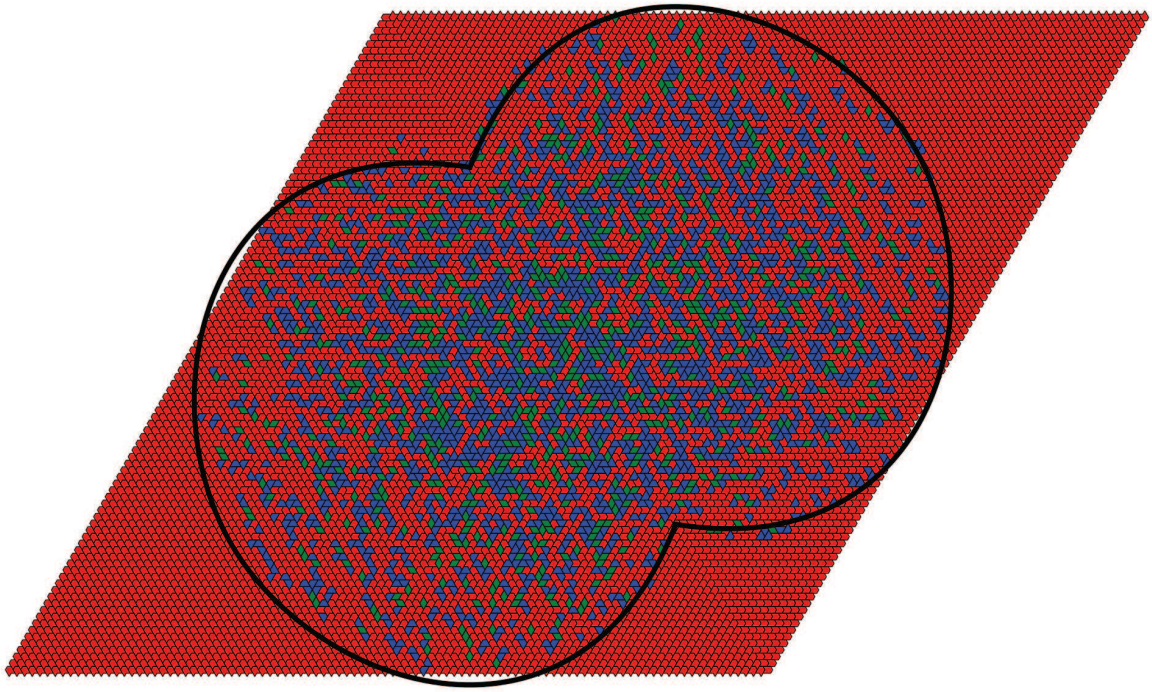

FIGURE 3.18: Kagome tiling of a lozenge of size 100 with a hand-drawn frozen boundary.

Conclusion

We have considered several mixing problems for tilings by a pair squares and the Kagome tilings. In restrained cases we get polynomial bounds on the mixing time of the considered Markov chains. We also provide simulations that suggest the *Artic circle* kind of behaviour on regions with non-flat boundaries. Simulations that were done by M. Tassy suggest such behaviour for tilings by bars as well as for the ribbon tilings of diamond-type regions (see Figure 3.19 [2]). We are hoping that this will arise more interest for studying dynamics on tiling systems. To begin with, ribbon tilings and T-tetraminos come to mind. They were studied from the structural and combinatorial point of view in [60, 87]. There is also a height function and flip-connectivity of the configurational space. This lays out a great foundation for studying Glauber dynamics on the set of tilings. The analysis seems to be challenging as well, but one could hope to get some restrained results as the ones presented above. That would be the first step to having some formal explanations about the dynamics of different tiling systems.

[2] Courtesy of M. Tassy.

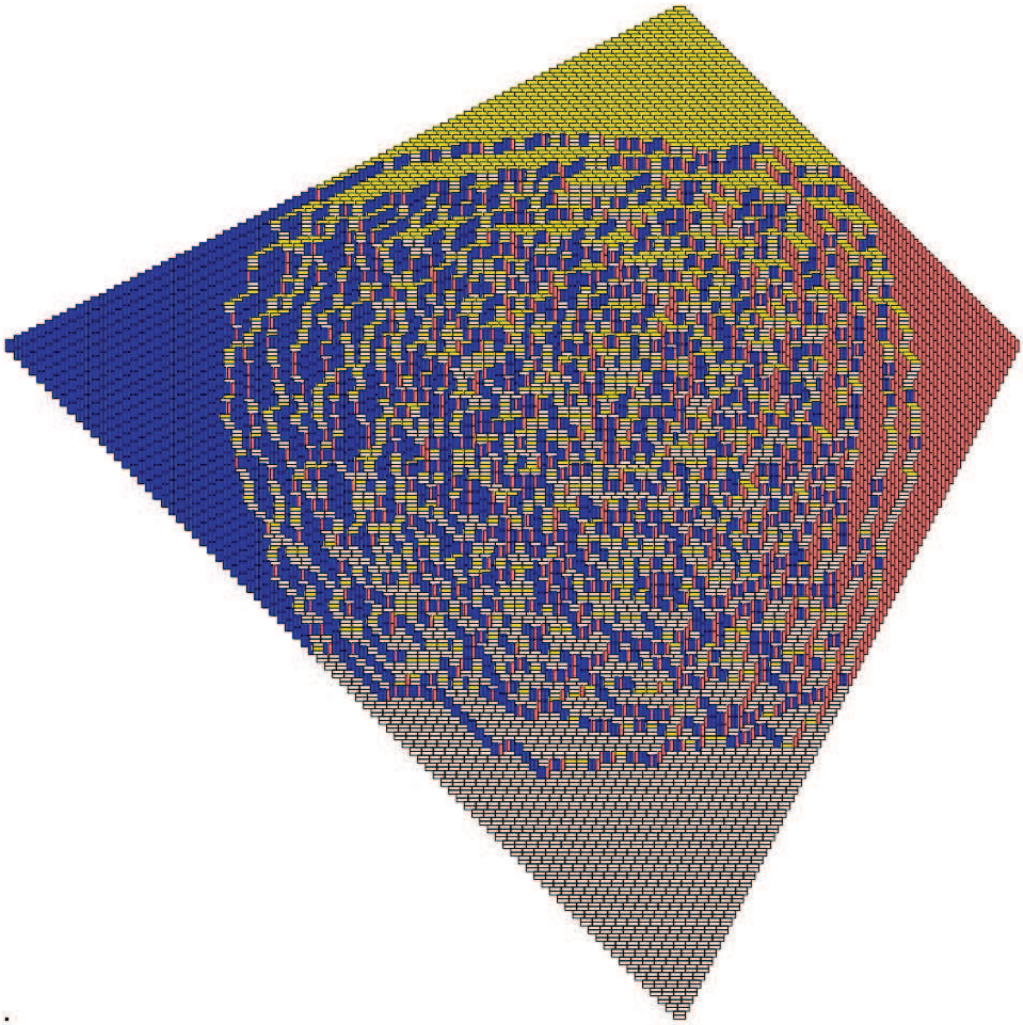

FIGURE 3.19: A tiling by 3×1 bars of a diamond-shaped region.

Appendices

Appendix A

Markov chains and mixing time

The following sections recalls classic definitions and theorems about Markov chains (see [1, 2, 37, 53, 63, 74]).

A.1 Markov chain

We will be dealing a lot with discrete random stochastic processes $(X_n)_{n\in\mathbb{N}}$ on a finite state space Ω that have the **memoryless** or **Markov property** which implies that the conditional distribution of X_n given $(X_{n-1}, \ldots, X_0)$ only depends on X_{n-1}. Normally we will be considering that the process is **homogeneous in time**: this means that the conditional distribution of X_n given that X_{n-1} equals some $x \in \Omega$, stays the same for all n. More formally, the definition is the following:

Definition A.1. A stochastic process $(X_n)_{n\in\mathbb{N}}$ on a finite state space $\Omega = \{x_1, \ldots, x_k\}$ is a **homogeneous Markov chain** if for all $n \in \mathbb{N}$ and
$i, j, i_0, \ldots, i_{n-2} \in \{1, \ldots, k\}$:

$$\mathbb{P}(X_n = x_j \mid X_{n-1} = x_i, \ldots, X_1 = x_{i_1}, X_0 = x_{i_0}) = \mathbb{P}(X_n = x_j \mid X_{n-1} = x_i). \tag{A.1}$$

The $k \times k$ matrix P with elements $P_{i,j} = \mathbb{P}(X_n = x_j \mid X_{n-1} = x_i)$, $i, j \in \{1, \ldots, k\}$ is its **transition matrix**.

Each element $P_{i,j}$ of the transition matrix P is the conditional probability to get from the state x_i to x_j. We shall sometimes write $P(x_i, x_j)$ for $P_{i,j}$. Every transition matrix satisfies

$$P_{i,j} \geq 0 \;\; \forall i, j \in \{1, \ldots, k\} \tag{A.2}$$

and

$$\sum_{j=1}^{k} P_{i,j} = 1 \ \ \forall i \in \{1, \ldots, k\}. \tag{A.3}$$

It is not the case for any transition matrix that $P_{i,j} = P_{j,i}$ for any $i, j \in \{1, \ldots, k\}$. When it is the case, the Markov chain M with this transition matrix is called **reversible**. A **period** p of a state $x \in \Omega$ is

$$p = gcdt > 0 : P^t(x, x) > 0 \tag{A.4}$$

If $p = 1$, the state x is called **aperiodic**. If every state $x \in \Omega$ is aperidoc (this means that there is a positive probability to stay in each state), the chain is called **aperiodic**. M is **irreducible** if $P(x_i, x_j) > 0$ for any two different states $x_i, x_j \in \Omega$.

Definition A.2. A Markov chain is said to be **ergodic** if it is aperiodic and irreducible.

An **initial distribution** for the Markov chain $(X_0, X_1, \ldots)$ is the following row vector $\mu^{(0)}(k)$ of size length k:

$$\mu^{(0)}(k) = (\mu_1^{(0)}, \ldots, \mu_k^{(0)}) = (\mathbb{P}(X_0 = x_1), \mathbb{P}(X_0 = x_2), \ldots, \mathbb{P}(X_0 = x_k)). \tag{A.5}$$

To get $\mu_{(1)}(k)$ one has to multiply $\mu_{(0)}(k)$ by the transition matrix P:

$$\mu^{(1)}(k) = \mu^{(0)}(k)P. \tag{A.6}$$

$\mu^{(n)}(k)$ can be obtained as follows:

$$\mu^{(n)}(k) = \mu^{(0)}(k)P^n. \tag{A.7}$$

A distribution μ is called **stationary** for the Markov chain with transition matrix P if $\mu P = \mu$.

Definition A.3. For two probability distributions ν and μ on $\Omega = \{x_1, \ldots, x_k\}$ the **total variation distance** between them is

$$d_{TV}(\nu, \mu) := \frac{1}{2} \sum_{i=1}^{k} |\nu_i - \mu_i|. \tag{A.8}$$

If there is a sequence $\{\mu^{(n)}\}_n \in \mathbb{N}$ of probability distributions on Ω, then it **converges to ν in total variation (TV)** as $n \longrightarrow \infty$ if

$$\lim_{n \to \infty} d_{TV}(\nu^{(n)}, \nu) = 0. \tag{A.9}$$

The following theorems are essential in the theory of Markov chains.

Theorem A.4 (Convergence theorem)**.** *Let $(X_0, X_1, \ldots)$ be an ergodic Markov chain on a finite state space Ω with transition matrix P and a random initial distribution $\mu^{(0)}$. Let μ be its stationary distribution. Then $\mu^{(n)}$ converges to μ in TV.*

Theorem A.5. *Any ergodic Markov chain has a unique stationary distribution.*

If a Markov chain is not only ergodic but also has a symmetric transition matrix, then its stationary distribution is **uniform**.

It is rather standard to denote the stationary distribution for an ergodic Markov chain by π (even though it might create confusion with the notation for the ratio of a circle's circumference to its diameter).

A.2 Mixing and Coupling times

We consider reversible ergodic Markov chains with a finite state space Ω. We denote its stationary distribution by π, its probability law by P. For any initial state $x \in \Omega$ let the total variation distance between $P(x, \dot{)}$ and π is

$$d_{TV}(P(x, \cdot), \pi) = \frac{1}{2} \sum_{y \in \Omega} |P^t(x, y) - \pi(y)|. \tag{A.10}$$

Let us write it as $d_x(t)$. The **mixing time** of the MC is the time it takes the chain to get close to its stationary distribution. Formally, it is defined as follows:

$$\tau_{mix}(\varepsilon) = \max_{x \in \Omega} \min\{t : d_x(t') \leq \varepsilon \; \forall \; t' \geq t\}, \tag{A.11}$$

$$\tau_{mix} := \tau_{mix}\left(\frac{1}{4}\right). \tag{A.12}$$

A classical way to bound the rate of convergence of a chain is to bound its mixing time. There are lot of different ways of bounding the mixing time: via the second largest eigenvalue (which can be analyzed using the corresponding tiling graph's properties), coupling methods (see, e.g., [9, 10, 22, 24, 25, 30, 51, 63, 71, 80, 89]).

Here we concentrate on the **coupling** method. A coupling for two probability distributions μ and ν is a pair of random variables (X, Y) defined on the same probability space

such that $\mathbb{P}(X = x) = \mu(x)$ and $\mathbb{P}(Y = y) = \nu(y)$. Here we will be using couplings for Markov chains where constructing copies of the chain proves to be a useful tool to analyze the distance to stationarity. A **coupling of a MC** is a stochastic process $(X_t, Y_t)_t$ on $\Omega \times \Omega$ such that:

1. X_t and Y_t are copies of the MC with initial states $X_0 = x$ and $Y_0 = y$;
2. If $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$.

Let $T^{x,y} = \min\{t : X_t = Y_t | X_0 = x, Y_0 = y\}$. Then define the **coupling time** of the MC to be

$$\tau_{cp} = \max_{x,y} \mathbb{E}T^{x,y}. \tag{A.13}$$

The following result [1] relates the coupling and mixing times:

Theorem A.6 (Aldous)**.**

$$\tau_{mix}(\varepsilon) \leq \tau_{cp}\lceil \ln \varepsilon^{-1} \rceil. \tag{A.14}$$

One of the most used methods to bound mixing time is the following path coupling theorem [24]. The authors show that in order to bound the coupling time, one only has to consider pairs of configurations of the coupled chain that are close to each other in the defined metric. It is sufficient to prove that they [each pair of configurations] have more tendency to remain close to each other under the evaluation of the chain. The mixing time then is polynomial and depends on the diameter of the corresponding graph.

Theorem A.7 (Dyer-Greenhill)**.** *Let $\varphi : \Omega \times \Omega \to \{0, \ldots, D\}$ be an integer-valued metric, U – a subset of $\Omega \times \Omega$ such that for all $(X_t, Y_t) \in \Omega \times \Omega$ there exists a path between them: $X_t = Z_0, Z_1, \ldots, Z_n = Y_t$ with $(Z_i, Z_{i+1}) \in U$ for $0 \leq i \leq r - 1$ and*

$$\sum_{i=0}^{r-1} \varphi(Z_i, Z_{i+1}) = \varphi(X_t, Y_t).$$

Let MC be a Markov chain on Ω with transition matrix P. Consider a random function $f : \Omega \to \Omega$ such that $\mathbb{P}(f(X) = Y) = P(X, Y)$ for all $X, Y \in \Omega$, and let a coupling be defined by $(X_t, Y_t) \to (X_{t+1}, Y_{t+1}) = (f(X_t), f(Y_t))$.

1. *If there exists $\beta < 1$ such that $\mathbb{E}[\varphi(X_{t+1}, Y_{t+1})] \leq \beta\varphi(X_t, Y_t)$ for all $(X_t, Y_t) \in U$, then the mixing time satisfies*

$$\tau_{mix}(\varepsilon) \leq \frac{\ln(D\varepsilon^{-1})}{1 - \beta}.$$

2. *If* $\beta = 1$, *let* $\alpha > 0$ *satisfy* $\mathbb{P}(\varphi(X_{t+1}, Y_{t+1}) \neq \varphi(X_t, Y_t)) \geq \alpha$ *for all* t *such that* $X_t \neq Y_t$. *Then the mixing time satisfies*

$$\tau_{mix}(\varepsilon) \leq \left\lceil \frac{eD^2}{\alpha} \right\rceil \lceil \ln \varepsilon^{-1} \rceil.$$

Appendix B

Random generation

This section recalls algorithms for random generation using Markov chains (some reference are: [20, 41, 44, 79, 90]). Given a finite state space Ω, a uniform **random sample** is a random selection from Ω, meaning that each element has the probability $\frac{1}{|\Omega|}$ to be selected. One can consider different systems, such as matchings, independent sets, tilings, etc (see, e.g,[23, 48, 59, 67]) Having a sample from uniform distribution allows to get an idea about the general look of the considered system, study its characteristics via simulations and approximate the size of the state space.

There is a general way to sample using a Markov chain $(X)_t$ that has a uniform distribution on Ω. When t is sufficiently large, X_t is approximately uniformly distributed (close enough with respect to some ε). The method of sampling from a given probability distribution is known as **Markov chain Monte Carlo**. So the general question is to estimate how large t must be so that X_t is sufficiently close from the desired distribution. If one has polynomial bounds on the mixing time of the chain, one can run an **fully polynomial almost uniform sampler (fpaus)**, which generates solutions from some distribution μ that is close to the uniform distribution π of our chain in total variance:

$$d_{TV}(\mu, \pi) \leq \delta \tag{B.1}$$

in polynomial time of the input size and $\log \delta - 1$ for some positive error δ.

B.1 Approximate counting

A general counting problem (e.g., counting the number of perfect matchings in a graph, the number of tilings of a region or computing the permanent of a matrix) can be seen as

computing a function $f : A^* \longrightarrow \mathbb{N}$, where A is a finite alphabet used to encode problem instances (e.g., the input matrix whose we would like to compute the permanent, or the number of tiles that are used for tiling the region). In general, counting problems are hard (belong to the $\#P$ class [1]). The goal becomes to get an approximation on the output of the counting problem. The algorithm used for this task is called a **fully polynomial randomized approximation scheme (fpras)**. Given an input $ax \in A^*$, parameter $\varepsilon, \delta > 0$ it computes an output Y such that

$$\mathbb{P}[(1-\varepsilon)f(x) \leq Y \leq (1+\varepsilon)f(x)] \geq 1-\delta \tag{B.2}$$

in time $t = poly(|x|, \varepsilon^{-1}, \log(1/\delta))$. Usually δ is taken to be $1/4$, so the algorithm outputs a good approximation (within ε from $f(x)$) with probability at least $3/4$.

If a problem is **self-reducible**, meaning that the solution set S can be expressed in terms of a polynomially bounded number of solution sets S_i such that S_i belongs to a smaller problem instance, the fpras is equivalent to fpaus [48]. This means that one has to run the almost uniform sampler for each S_i an enough number of times. And put it all together to get the approximation for the total S. The problem is always that when one wants to get a sample of S_i, it usually means that one has to run the corresponding Markov chain long enough for it to be close enough to its stationary distribution. It means that one has to have a good polynomial upper bound on its relaxation time or mixing time, otherwise sampling will be computationally untrackable. Generally that is a challenging question (as we have seen throughout Chapter 1–3).

B.2 Grand coupling or sanwiching

Consider a grand coupling of a given MC, which constructs $|\Omega|$ marginal copies $(X_t^1, \dots X_t^{|\Omega|})_t$ of the initial chain $(X_t)_t$ and runs them simultaneously according to the same rules at each step. Once the algorithm reaches the step t_e such that all $X_{t_e}^i$ are equal, they behave like a single chain $(X_t)_t$. If the chain is monotone then for a coupling $(X_t, Y_t)_t$ on $\Omega \times \Omega$ and $A, B \in \Omega$ and some $t > 0$ if $X_t = A \leq_\varphi Y_t = B$ according to some metric φ defined on Ω, then $X_{t+1}(A) \leq_\varphi Y_{t+1}(B)$. Let $A_{max}, A_{min} \in \Omega$ be the maximal and minimal configurations according to φ (suppose they are unique, otherwise there will just be a set of maximal and minimal tilings). Start the coupling $(X_t, Y_t)_t$ at $X_0 = A_{max}$, $Y_0 = A_{min}$. Monotonicity implies that once X_t and Y_t meet, all states are coupled. So there is no need of doing the grand coupling in all states simultaneously.

[1] $\#P$ class contains the set of counting problems associated with decision problems that belong to the class NP.

B.3 Coupling from the Past

Standard methods of getting a sample allow to get a sample not from the stationary distribution π, but from the distribution μ that is close π. It can be as close as we want to, but the sample we get will still not be completely uniform. There is a coupling technique called **Coupling from the Past** (see [36, 63, 79]), using which one can get an exact sample from π. As the name suggests, the coupling is being done backwards, by starting at some time $-t$, $t > 0$. If by the time 0 the chains (which are copies of the initial chain $(X_t)_t$) have not coalesced, go to the time $-2t$ and proceed using the same transformations on the stage from $-t$ to 0. Let us denote the evolution of the chain from time t_1 to t_2 as follows:

$$F_{t_1}^{t_2} = (f_{t_2-1} \circ f_{t_2-2} \circ \ldots \circ f_{t_1})(x), \tag{B.3}$$

where $X_{t+1} = f(X_t, Z_t)$, f defines the coupling and Z_t is drawn u.a.r from $[0, 1]$ at each time t.

An advantage if this algorithm is that it allows to sample from π by only going a finite number of steps backwards (as opposed to coupling in the futur). Suppose we found T s.t. $F_{-T}^0(x) = F_{-T}^0(y)$ for all states x, y from the state space. Take $t > T$, then

$$F_{-t}^0(x) = F_{-T}^0 \circ f_{-T-1} \circ \ldots \circ f_{-t}(x) = F_{-T}^0(z) \tag{B.4}$$

for some state z. But $F_{-T}^0(z)$ just equals $F_{-T}^0(x)$, so instead of going to $-\infty$ to get a sample from π it is sufficient to only go to $-T$.

If the chain is monotone, coupling from the past becomes very efficient to implement, because one can consider a coupling that starts the minimal and maximal states instead of the grand coupling.

Bibliography

[1] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, Seminaire de Probabilites 1981/1982, Springer Lecture Notes in Mathematics **986**, pp. 243–297.

[2] D. Aldous, *Some inequalities for reversible Markov chains*, J. London Math Society **25**, 1982, pp. 564–576.

[3] R.J. Baxter, *Hard hexagons: Exact solution*, J. Physics A **13**, 1980, pp. 1023–1030.

[4] R.J. Baxter, *Exactly solved models in statistical mechanics*, Academic Press, London, 1982.

[5] A. Berger, S. Rechner, *Broder's chain is not rapidly mixing*, arXiv:1404.4240v1 [cs.DM], 2014.

[6] O. Bodini, *Tilings on the butterfly lattice*, European Journal of Combinatorics **27**, 2006, pp. 1082–1087.

[7] O. Bodini, Th. Fernique, D. Regnault, *Stochastic flips on two-letter words*, Analytic Algorithmics & Combinatorics, 2010, pp. 48–55.

[8] O. Bodini, Th. Fernique, E. Rémila, *A Characterization of flip-accessibility for rhombus tilings of the whole plane*, Inf. Comput. **206**, 2008, pp. 1065–1073.

[9] M. Bordewich, M. Dyer, *Path coupling without contraction*, Preprint submitted to Elsevier Science, 2005.

[10] R. Bubley, M. Dyer, *Path coupling: a technique for proving rapid mixing in Markov chains*, 38th Annual Symposium on Foundations of Computer Science, 1997, pp. 223–231.

[11] S. Cannon, S. Miracle and D. Randall, *Phase transitions in random dyadic tilings and rectangular dissections*, 26th Symposium on Discrete Algorithms (SODA), 2015.

[12] N.J. Calkin, K. James , S. Purvis, S. Race, K. Schneider, M. Yancey, *Counting kings: as easy as* $\lambda_1, \lambda_2, \lambda_3$..., Congr. Numer **183**, 2006, pp. 83–95.

[13] P. Caputo, F. Martinelli, A. Sinclair, adn A. Stauffer, *Random lattice triangulations: Structure and algorithms*, Annals of Applied Probability **25**, 2015, pp. 1650–1685.

[14] P. Caputo, F. Martinelli, F.L. Toninelli , *Mixing Times of Monotone Surfaces and SOS Interfaces: A Mean Curvature Approach*, Communications in Mathematical Physics **311** (Issue 1), 2012, pp. 157–189.

[15] H. Cohn, N. Elkies, J. Propp, *Local statistic for random domino tilings of the Aztec diamond*. Duke Math, J. **85**, 1996, pp. 117–166.

[16] H. Cohn, R. Kenyon, J. Propp, *A variational principle for domino tilings*, J. American Mathematical Society **14**, 2001, pp. 297–346.

[17] H. Cohn, M. Larsen, J. Propp, *The shape of a typical boxed plane partition*, New York Journal of Mathematics **4**, 1998, pp. 137–165.

[18] J. Conway, J. Lagarias, *Tilings with polyominos and combinatorial group theory*, Journal of Combinatorial Theory A **53**, 1990, pp. 183–208.

[19] N. Destainville, *Entropy and boundary conditions in random rhombus tilings*, J. Phys. A: Math. Gen. **31**, 1998, pp. 6123–6139.

[20] P. Diaconis, S. Holmes, R.M. Neal, *Analysis of a nonreversible Markov chain sampler*, Annals of Applied Probability **10**, 2000, pp. 726–752.

[21] P. Diaconis, L. Saloff-Coste, *Comparison theorems for reversible Markov chains*, Annals of Applied Probability **3**, 1993, pp. 696–730.

[22] P. Diaconis, D. Stroock, *Geometric bounds for eigenvalues of Markov chains*, Annals of Applied Probability **1**, 1991, pp. 36–61.

[23] M. Dyer, A. Frieze, M. Jerrum, *On counting independent sets in sparse graphs*, SIAM J. Computing **33**, 2002, pp. 1527–1541.

[24] M. Dyer, C. Greenhill, *A More Rapidly Mixing Markov Chain for Graph Colorings*, Random Structures and Algorithms **13**, 1998, pp. 285–317.

[25] M. Dyer, C. Greenhill, *On Markov chains for independent sets*, Journal of Algorithms **35**(Issue **1**), 2000, pp. 17–49.

[26] N. Elkies, G. Kuperberg, M. Larsen , and J. Propp, *Alternating sign matrices and domino tilings*, J. Algebraic Combin. **1**, 1992 , pp. 111–132 and 219–234.

[27] V. Elser, *Solution of the dimer problem on a hexagonal lattice with boundary*, J. Phys. A: Math. Gen. **17**, 1984, pp.1509–1513.

[28] V. Elser, *Comment on quasicrystals: a new class of ordered structures*, Phys. Rev. Lett **54**, 1985.

[29] Th. Fernique, D. Regnaut, *Stochastic Flips on Dimer Tilings*, DMTCS proc. AM, 2010, pp. 205–218.

[30] J.A. Fill, *Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with application tothe exclusion process*, Annals of Applied Probability **1**, 1991, pp. 62–87.

[31] Ph. Flajolet, S. Sedgewick, *Analytical Combinatorics*, Cambridge University Press, 2009.

[32] M. Garey and D. S. Johnson, *Computers and intractability: a guide to the theory of NP-completeness*, Freeman, San Francisco, 1979.

[33] I. Gessel, G. Viennot, *Binomial Determinants, Paths and Hook Length Formulae*, Advances in Mathematics **58**, 1985, pp. 300–321.

[34] J. de Gier, B. Nienhuis, *Exact Solution of an Octagonal Random Tiling Model*, Journal of Statistical Physics **87** (Issue 1), 1997, pp. 415–437.

[35] B. Grünbaum and G.C. Shephard, *Tilings and Patterns*, W.H. Freeman, New York, 1987.

[36] O. Häggström, K. Nelander, *On exact simulation from Markov random fields using coupling from the past*, Scand. J. Statist. **26(3)**, 1999, pp. 395–411.

[37] O. Häggström, *Finite Markov Chains and Algorithmic Applications*, London Mathematical Society, Student Texts **52**, Cambridge University Press, 2002.

[38] O.J. Heilmann, E.H. Lieb, *Theory of monomer-dimer systems*, Comm. in Mathematical Physics **25**, 1972, pp. 190–232.

[39] S. Heubach, *Tiling an m-by-n Area with Squares of Size up to k-by-k with $m \leq 5$*, Congressus Numerantium **140**, 1999, pp. 43–64.

[40] S. Heubach, P. Chinn, R. Grimaldi, *Patterns Arising From Tiling Rectangles with 1x1 and 2x2 Squares*, Congressus Numerantium **150**, 2000, pp. 173–192.

[41] M. Huber, *Exact sampling and approximate counting techniques*, in: Proceedings of the 30th Symposium on the Theory of Computing, 1998, pp. 31–40.

[42] C. Janot, *Quasicrystals: The state of the art*, Oxford University Press, 1992.

[43] M. Jerrum, *Two-dimensional monomer-dimer systems are computationally intractable*, Journal of Statistical Physics **48**, 1987, pp. 121–134.

[44] M.R. Jerrum, *Counting, Sampling and Integrating: Algorithms and Complexity*, Birkhäuser Verlag, Basel, 2003.

[45] M.R. Jerrum, A.J. Sinclair, *Approximating the permanent*, SIAM Journal on Computing **18**, 1989, pp. 1149–1178.

[46] M. Jerrum and A. Sinclair, *The Markov chain Monte Carlo method: an approach to approximate counting and integration*, in D. Hochbaum, ed., Approximation Algorithms or NP-Hard Problems, PWS Publishing, Boston, 1996, pp. 482–520.

[47] M. Jerrum, A. Sinclair and E. Vigoda, *A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries*, Journal of the ACM, Vol. **51**, no. 4, 2004, pp.671–697.

[48] M. Jerrum, L. Valiant, V. Vazirani, *Random generation of combinatorial structures from a uniform distribution*, Theoret. Comput. Sci. **43**, 1986, pp. 169–188.

[49] W. Jockush, J. Propp, P. Shor, *Random Domino Tilings and the Arctic Circle Theorem*, arXiv:math/9801068v1 [math.CO], 1998.

[50] P.A. Kalugin, *The square-triangle random-tiling model in the thermodynamic limit*, J. Phys. A: Math. Gen. **27**, 1994, pp. 35–99.

[51] A. Karzanov, L. Khachiyan, *On the conductance of order Markov chains*, Order **8(1)**, 1991, pp. 7–15.

[52] P.W. Kasteleyn, *The statistics of dimers on a lattice, I. The number of dimer arrangements on a quadratic lattice*, Physica **27**, 1961, 1209–1225.

[53] J. Keilson, *Markov Chain Models: Rarity and Exponentiality*, Springer-Verlag, New York, 1979.

[54] C. Kenyon, R. Kenyon, *Tiling a Polygon with Rectangles*, Proc. 33rd FOCS, 1992, pp. 610–619.

[55] R. Kenyon, *A Note on Tiling with Integer-sided Rectangles*, Journal of Combinatorial Theory **74** (Issue 2), 1996, pp. 321–332.

[56] R. Kenyon, *Local Statistics of Lattice Dimers*, Ann. Inst. H. Poincaré, Prob. et Stat. **33**, 1997, pp. 591–618.

[57] R. Kenyon, *An Introduction to the Dimer Model*, Lectures given at the School and Conference on Probability Theory, Trieste, 13-31 May, 2002.

[58] R. Kenyon, A. Okounkov, S. Sheffield, *Dimers and Amoebae*, Annals of Mathematics **163(3)**, 2003.

[59] C. Kenyon, D. Randall, A. Sinclair, *Approximating the number of dimer coverings of a lattice*, Journal of Statistical Physics **83**, 1996, pp. 637–659.

[60] M. Korn, I. Pak, *Tilings of rectangles with $T-$tetraminos*, Theoret. Comput. Sci. **319**, 2004, pp. 3–27.

[61] M.-L. Lackner, M.Wallner, *An invitation to analytic combinatorics and lattice path counting*, lecture notes, ALEA in Europe Young Researchers' Workshop, Bath, UK, 2015

[62] B. Laslier, F. L. Toninelli, *Lozenge tilings, Glauber dynamics and macroscopic shape*, Comm. Math. Phys. **338**, 2015, pp. 1287–1326.

[63] D.A. Levin, Y. Peres, E.L. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, 2009.

[64] E.H. Lieb, *Exact solution of the problem of the entropy of two-dimensional ice*, Phys. Rev. Lett. **18**, 1967, pp. 692–694.

[65] D. Lind, B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.

[66] M. Luby, D. Randall, A. Sinclair, *Markov Chain Algorithms for Planar Lattice Structures*, Proceedings in the 36th IEEE Symposium on Foundations of Computer Science, 1995, pp. 150–159.

[67] M. Luby, E. Vigoda, *Fast Convergence of the Glauber Dynamics for Sampling Independent Sets: Part I*, Random Structures & Algorithms, Volume **15**, Issue **3-4**, 1999, pp. 229–241.

[68] P. A. MacMahon, *Combinatory Analysis*, Cambridge University Press, 1915–16 (reprinted by Chelsea Publishing Company, New York, 1960).

[69] R.J. Mathar, *Tilings of Rectangular Regions by Rectangular Tiles: Counts Derived from Transfer Matrices*, arXiv:1406.7788 [math.CO], 2014.

[70] J. Matthews, *Markov chains for sampling matchings*, PhD thesis, Edinburgh University, 2008.

[71] M. Mihail, *Combinatorial Aspects of Expanders*, PhD Thesis, Harvard University, 1989.

[72] J. Nilsson, *On Counting the Number of Tiling of a Rectangle with Squares of Size 1 and 2*, 2016, 16 pages, preprint.

[73] E. Noradenstam, B. Young, *Domino shuffling on Novak half-hexagons and Aztec half-diamonds*, Elect. J. of Combin. **18**, 2011.

[74] J.R. Norris, *Markov Chains*, Cambridge University Press, 1997.

[75] I. Pak , *Tile invariants: new horizons*, Theoret. Comput. Sci. **303**, 2003, pp. 303–331.

[76] I. Pak, A. Sheffer, M. Tassy, *Fast Domino tileability*, arXiv:1507.00770, 2015.

[77] R. Pavlov, M. Schraudner, *Entropies realizable by block gluing $\mathbb{Z}^d$ shifts of finite type*, J. Analyse Mathématique, Volume **126**, Issue **1**, 2015, pp. 113–174.

[78] J. Propp, *Generalized domino-shuffling*, Theor. Comput. Sci., **303**(2-3), 2003, pp. 267–301.

[79] J. Propp, D. Wilson, *Exact Sampling with coupled Markov chains and applications to statistical mechanics*, Random Structures and Algorithms **9**, 1996, pp. 223–252.

[80] D. Randall, *Rapidly Mixing Markov Chains with Applications in Computer Science and Physics*, IEEE CS and the AIP Computing in Science & Engineering, 2006, pp. 30-41.

[81] D. Randall, P. Tetali, *Analysing Glauber dynamics by comparison of Markov Chains*,Journal of Mathematical Physics **41**, 2000, pp. 1598–1615.

[82] E. Rémila, *Tiling a polygon with two kinds of rectangles*, Algorithms – ESA 2004, V. 3221 of the series Lecture Notes in Computer Science, pp. 568–579.

[83] J.M. Robson, *Sur le pavage de figures du plan par des barres*, Actes des Journées Polyominos et Pavages, 1991, pp. 95–103.

[84] R. Robinson, *Undecidability and Nonperiodicity for Tilings of the Plane*, Inventiones Mathematicae, **12(3)**, 1971.

[85] N. Rolin, A. Ugolnikova, *Tilings by* 1×1 *and* 2×2, RAIRO-Theor. Inf. Appl. **50(1)**, 2016, pp. 105–116.

[86] D. Romik, *Arctic circles, domino tilings and square Young matrices*, Annals of Probability, **40**, 2012, pp. 611–647.

[87] S. Scheffield, *Ribbon tilings and multidimensional height functions*, Trans. of AMS **354**, 2002, pp. 4789–4813.

[88] M. Senechal, *Quasicrystals and geometry*, Cambridge University Press, 1995.

[89] A. Sinclair, *Improved bounds for mixing rates of Markov chains and multi-commodity flow*, Combinatorics, Probability & Computing **1**, 1992, pp. 351–370.

[90] A. Sinclair, *Counting and Generating Combinatorial Structures: a Markov Chain approach*, Birkhäuser, Boston, 1993.

[91] N.J.A. Sloane, and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press Inc., 1995.

[92] J.M. Stembridge, *Non-intersecting paths, Pfaffians, and plane partitions*, Adv. Math. **83**, 1990, pp. 96–131.

[93] H. N. V. Temperley and M. E. Fisher, *Dimer problem in statistical mechanics – an exact result*, Phil. Mag. **6**, 1961, 1061–1063.

[94] W.P. Thurston, *Conway's tiling groups*, The American Mathematical Monthly **97 (8)**, 1990, pp. 757–773.

[95] B. de Tilière *Partition function of periodic isoradial dimer models,* Probability Theory and Related Fields **138**, 2007, pp. 451–462.

[96] B. de Tilière *Quadri-tilings of the Plane*, Probability Theory and Related Fields **138**, 2007, pp. 487–518.

[97] L.G. Valiant, *The complexity of computing the permanent*, Theoretical Computer Science **8**, 1979, pp. 189–201.

[98] F. Wung, F.Y. Wu, *Exact solution of close-packed dimers on the kagome lattice*, Phys. Rev. E **75**, 2007.

[99] D.B. Wilson, *Mixing Times of Lozenge Tiling and Card Shuffling Markov Chains*, The Annals of Applied Probability **14**, 2004, pp. 274–325.