On the mixing time of the face flip– and up/down Markov chain for some families of graphs.

vorgelegt von
Diplom–Mathematiker
Daniel Heldt
geb. in Duisburg

Von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
– Dr.rer.nat. –

genehmigte Dissertation

Promotionsausschuss:
Vorsitzender: Prof. Dr. Jörg Liesen
Gutachter: Prof. Dr. Stefan Felsner
Gutachter: Prof. Dr. Günter Rote

Tag der wissenschaftlichen Aussprache: 18. Juni 2015

Berlin 2016

Preface

Creating this work I had some help:

Chapter 2 is joint work with Stefan Felsner and Peter Winkler. So far it was only presented in the Seminar *Anwendungen von Zufall und Wahrscheinlichkeit auf Fragen der Diskreten Mathematik* of Stefan Felsner in a presentation of Laura Vargas Koch in July 2014. This presentation was based on a preliminary version of chapter 2. I thank Laura Vargas Koch for some valuable remarks to this version, which where integrated into the chapter in this document. In addition we want to thank Benjamin Armbruster for the brief email communication regarding his version of Strassen's theorem on stochastic domination and Hendrik Lüthen for the implementation of some shapes to check, whether or not they suffice to find a bound for the mixing time of the up/down Markov chain M with the machinery presented in this work.

Chapter 3 is conjoint work with Stefan Felsner. We presented a preliminary version of this work at the 7th International Conference on Lattice Path Combinatorics and Applications [1] and got some feedback, which we really appreciate. We like to thank for literature hints, remarks and further ideas, which partly were included here. In particular, we thank Cyril Banderier, Christian Krattenthaler, and Helmut Prodinger for their remarks.
Additionally, a version containing the same content as this chapter was accepted for publishing by the Journal of Integer Sequences as [23] in December 2014.

Chapter 4 started as a simplification of an family of counter examples, which were presented in a preliminary version of [39]. Together with Stefan Felsner, we extended this to the content, which is presented in this thesis. So far, only the authors of [39] received an early preprint of this chapter.

As already pointed out, I owe advisor Stefan Felsner quite some gratitude. He advised me well. He pointed out the beauty of combinatorics to me and taught me some "Wiener Gelassenheit", when it comes to organize teaching. Thank you! In addition, he created the Arbeitsgruppe, I was member of for more then five years. Besides the people, I miss good riddles and good coffee most.

Besides the research, presented in this document, I was able to contribute to some research of Kolja Knauer and Torsten Ueckerdt, who both were members of Stefan Felsner's group at the Technical University of Berlin. This research on *Bend Numbers of Graphs* led to the publications [30] and [31], which was followed up by my two coauthors with [35]. I thank both for their kindness and collaboration.

Last but not least I thank my children, Frederic Alexander, Felix Anton, and Lisa Katharina for existence, gorgeous times and sleepless nights as well as my wife, Corinna Melanie Heldt: Möge es noch lange dauern, bis Deine guten Zeiten beginnen.

Daniel Heldt
Dortmund, December 2014

Contents

Introduction

This body of work engages in two natural Markov chains, each living on a family of finite distributive lattices and some somewhat related results on the enumeration of lattice paths.

The interest in these Markov chains rose due to their state spaces, the sets they are living on. Basic open questions for these sets are as simple as: *How big is the set?*, *How big is it asymptotically?*, or *How does a typical element in the set look like?*. While being finite, these sets are quickly to big to write down explicitly, so some machinery is required to give at least some answers to this questions to some extend.

In [25] Felsner and Zickfeld considered the first two of these questions for α-orientations and gave some bounds to the number of α-orientations for some families of graphs. This work now continues somewhat their work and considers the third question or at least, works somewhat into the direction to be able to answer it for some families of graphs. To find a typical element of a set, *sampling* is an applicable method if you have a Markov chain on these set, which mixes fast enough. Therefore, considering Markov chains on these sets was an approach to work on these questions. Note, that with reasonable *nice* Markov chains, one could even gather bounds on the size of their state spaces, as stated in Chapter 3.2 of [33].

Due to the structure of the sets, there are some natural Markov chains to consider, so the starting point for this work was the application of a coupling method – the *block coupling technique* – to the of α-orientations of some families of graphs to investigate the behaviour of the face flip Markov chain on these distributive lattices. An α-orientation of an undirected plane graph is an orientation of its edges, such that every vertex v has a by α prescribed outdegree $\alpha(v)$. The set of these orientations for a fixed α and graph forms a distributive lattice, whose elementary moves are flips, i.e. reversing the orientation of all edges of a directed cycle. For somewhat nice graphs and α, this can be further restricted to face flips. This means it suffices to reverse facial cycles. Now the face flip Markov chain somewhat uniformly at random reverses one of the oriented facial cycles and wanders this way through the distributive lattice of α-orientations. Chapter 2 covers the block coupling technique and its application to some families of distributive lattices. These lattices are the sets of k-heights of some families of planar graphs, mostly for $k = 2$. Here k-heights assign each face of a plane graph an integer number between 0 and k, such that faces, having an edge in common, bear numbers, which differ at most by 1. They are closely related to α-orientations (interpret the assigned numbers as indication, how often this face has to be flipped, starting from a fixed orientation), but the resulting lattices are more inoffensive towards the block coupling method, so we were able to apply it there.

The block coupling method relies on some enumeration, which we did with the help of computers. As it happens quite often in combinatorics, they quickly showed us their limits. Therefore this work wanders from the original subject to the enumeration of somewhat 1-*dimensional k-heights*, or bounded lattice paths. I hoped to understand the one dimensional behaviour and translate its essence back to the two-dimensional case of planar graphs. This digression forms Chapter 3. It contains some results, which are interesting on their own, but fail to contribute to the original goal, because I did not make the way back to k-heights. Still, this chapter contains explicit formulas to enumerate bounded Motzkin paths, which are the same as the set of k-heights of a sequence of adjacent faces. This enumeration then made itself independent and covered a lot of different integer sequences, of whom some looked quite independent beforehand. Also, it applies nicely some graph theory and linear algebra to enumerate lattice paths, which might be a method worth further investigation.

Finally, the fourth chapter, containing some original content covers α-orientations and their face flip Markov chain. Although Chapter 4 does not make use of the block coupling method, we were able to make some progress using tower moves. These moves are another speed up technique which can be applied on top of path coupling in the same manner as block coupling. Here, we were able to improve some known results: Besides other results, the chapter lays out a family of quite harmless graphs, such that the face flip Markov chain is not rapidly mixing the set of α-orientations of these graphs, which in the end answers the original question to quite some extent, though not in a positive way.
So, going back to the initial question, we were able to show, that sampling of α-orientations works for some very constrained families of graphs (except for some hilarious coefficients, which might be improved in the future substantially) and does not work for others. Still, on the way to these results, we will come across quite some other results.

1 Preliminaries

In the Introduction we already mentioned briefly the two main structures, which build the foundation of this work. They are α-orientations and their distributive lattices on the one hand and Markov chains and their mixing times on the other. We introduce both very briefly in the next two sections and give some more of the required background in situ in the following chapters:

1.1 α-Orientations

Let $G = (V, E)$ be a simple plane graph with vertex set V and edge set E. Further, let $\alpha : V \to \mathbb{N}$ be a weight function, living on the vertices of G.

Definition 1 (α-orientations)
An *α-orientention* $\vec{G}$ of G is an orientation of the edge-set of G, such that every vertex $v \in V$ has outdegree $\alpha(v)$. Denote the set of all α-orientations with $\Omega_{\alpha,G}$ or Ω for short, if G and α cannot be mistaken. We call α *feasible*, if there exists an α-orientation. Further, edges are called *rigid*, if they are oriented in the same way in all α-orientations $\Omega_{\alpha,G}$.

The notion of α-orientations was introduced by Felsner in [20] as a generalization of 3-orientations of planar triangulations, which are in bijection to Schnyder–woods. They are on the other hand also a generalization of 2-orientations of the rectangle grid (or plane quadrangulations). In the same paper Felsner showed, that for plane G and a feasible α (i.e. one, such that there are α-orientations of G) the set Ω forms a distributive lattice and described explicitly the lattice's edges as flips of the orientation of directed *essential cycles*. Here, these essential cycles are exactly the facial cycles of all faces which do not share an edge with the outer face. Again according to Felsner, the set of α-orientations is in bijection with the set of α-potentials. These potentials basically count the number of flips of every essential cycle, on a path from the lattice's minimum to the corresponding orientation.

Definition 2 (essential cycles and α-potentials)
More formally, a simple cycle C of G is an essential cycle if and only if

- C is cord-free
- the interior cut of C is rigid
- there exists an α-orientation $\vec{X}$ of G, such that C is oriented in $\vec{X}$

Further, let $\mathcal{C}$ be the set of essential cycles and $\vec{G}_{\min}$ the minimum of Ω, then $\phi : \mathcal{C} \to \mathbb{N}$ is an *α-potential* of G if

- $|\phi(C) - \phi(C')| \leq 1$ for all $C, C' \in \mathcal{C}$, which share an edge e

- $\phi(C) \leq 1$ if C is the only essential cycle containing an edge e
- if $C^{l(e)}$ and $C^{r(e)} \in \mathcal{C}$ are the cycles left and right of e in $\vec{X}_{\min}$, we have $\phi(C^{l(e)}) \leq \phi(C^{r(e)})$

In the considered cases the set of essential cycles corresponds to the set of faces (or to the cycles, defined by the faces), except the outer face and all faces sharing at least one edge to the outer face.

1.2 Markov chains

The Markov chains in the following chapters always have a finite state space and are discrete-time Markov chains. A nice introduction to this kind of Markov chains and some remarkable applications of them to combinatorics, can be found in Jerrum's book [33]. A more recent overview is the book [38]. Jerrum's definition of a Markov chain is as follows:

Definition 3 (Markov chain)
Let Ω be a finite state space. A sequence $(X_t \in \Omega)_{t=0}^{\infty}$ of random variables is a *Markov chain* on Ω if

$$\Pr[X_{t+1} = y \mid X_t = x_t, X_{t-1} = x_{t-1}, \dots, X_0 = x_0] = \Pr[X_{t+1} = y \mid X_t = x_t],$$

for all $t \in \mathbb{N}$ and $x_t, x_{t-1}, \dots, x_0 \in \Omega$. This means the next state only depends on the current and not on the previous ones. In addition, as Jerrum does, we always assume the Markov chain to be (time-) homogeneous, which means

$$\Pr[X_{t+1} = y \mid X_t = x] = \Pr[X_{t'+1} = y \mid X_{t'} = x]$$

for all $t, t' \in \mathbb{N}$. So a Markov chain is not aware of its history and has always the same transition probabilities.

This lack of history allows to capture the whole chain with a *transition matrix* $\mathcal{M} \in \mathbb{R}^{\Omega \times \Omega}$ (i.e. a matrix $\in \mathbb{R}^{|\Omega| \times |\Omega|}$ with entries indexed by Ω), where $\mathcal{M}_{x,y} = \Pr[X_{t+1} = y \mid X_t = x]$. So the entry of $\mathcal{M}$ associated with (x, y) is the transition probability from x to y (which does not change with time). Note, that now the multiplication of a stochastic row vector v with $\mathcal{M}$ corresponds with the application of one step of the Markov chain M to some random variable X distributed according to v (i.e. X_t is with probability v_y at y for all $y \in \Omega$), i.e. $v \cdot M$ is a stochastic row vector, describing the distribution of X_{t+1}.

Next, we can interpret $\mathcal{M}$ as adjacency matrix of a directed simple graph G_P (with loops), where $\mathcal{M}_{x,y}$ is the edge-weight of the edge from x to y. Now, a Markov chain can be easily understood as a random walk on the Graph G_P, where the wanderer on vertex x chooses his or her next step, according to the weights of the edges, leaving x. We continue with some properties, before we expose what we are up for:

Definition 4 (aperiodicity, irreducibility and stationary distributions of Markov chains)
Let M be a Markov chain on a finite state space Ω.

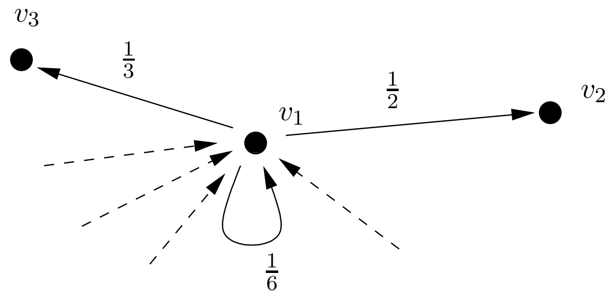

Figure 1.1: A wanderer at vertex v_1 goes to v_2 with probability $\frac{1}{2}$, to v_3 with probability $\frac{1}{3}$ and stays at v_1 otherwise.

- We say, M is aperiodic, if (some advance assumed), we are able to be at any element of the state space at any time (assuming strong connectivity). Formally, if $\gcd\{t \mid \mathcal{M}^t_{x,x} > 0\} = 1$ for all $x \in \Omega$.
- We say, M is *irreducible*, if there is a path of edges with positive weight from every $x \in \Omega$ to every $y \in \Omega$ (i.e. $G_{\mathcal{M}}$ is strongly connected). Formally, if for all $x, y \in \Omega$ there some $t \in \mathbb{N}$ such that $\mathcal{M}^t_{x,y} > 0$.
- If M is aperiodic and irreducible, we say M is *ergodic* (because Ω is finite).
- A distribution π on Ω is a stationary distribution of M if the distribution of some random variable X, distributed according to π does not change, when a step of M is applied to X, this means, π written as stochastic row vector is a left eigenvector with eigenvalue 1 of the transition matrix $\mathcal{M}$ of the Markov chain M.

The next well known theorem states, why these properties matter:

Theorem 1 *(Theorem 3.6 of [33])*
Let M be an ergodic Markov chain on finite state space Ω. Now, M has a unique stationary distribution π and

$$\mathcal{M}^t_{x,y} \longrightarrow \pi(y) \text{ for } t \longrightarrow \infty.$$

Note, that both, irreducibility and aperiodicity are required.
Figure 1.2 shows to Markov chains in their graph representation. The left on is not irreducible, since there is no path from vertex v_1 to v_2. The matrix $\mathcal{M}_{G_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ has an eigenspace of dimension two for the eigenvalue 1, so there is more than one stationary distribution. The stochastic vectors $(1, 0, 0)$ and $(0, \frac{1}{2}, \frac{1}{2})$ are both left eigenvectors with eigenvalue 1 of the transition matrix, so each of them corresponds to a stationary distribution. The right Markov chain, specified by G_2, is not aperiodic because every path from a vertex back to itself is of even length. Therefore, the second property of the theorem does not hold: If we start at a red vertex, we are always at a red vertex after an even number of steps and at a blue one, after an odd number of steps. Therefore, $\mathcal{M}^T_{x,y}$ does not converge.

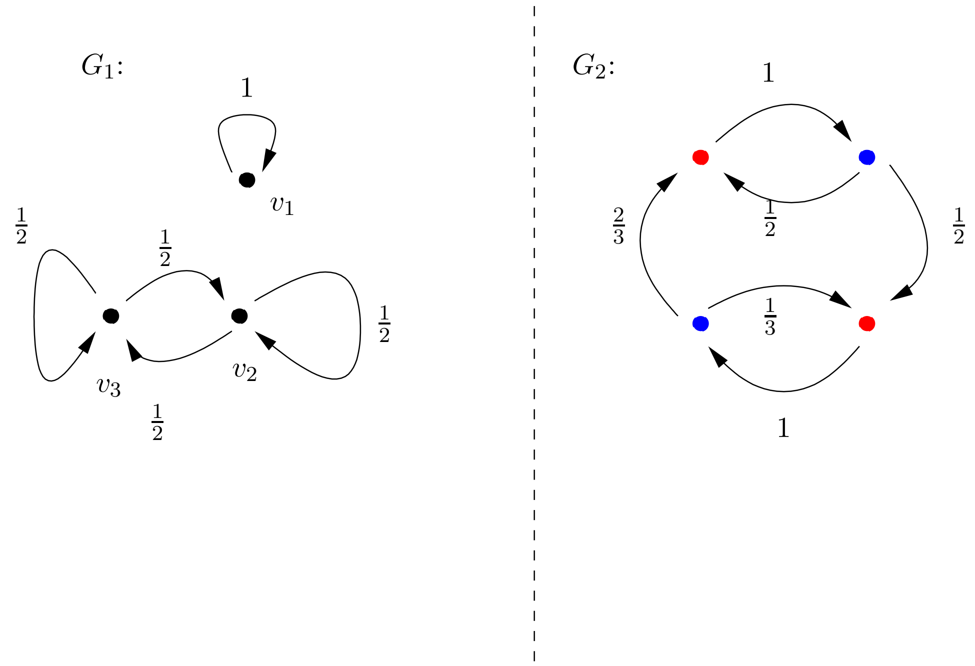

Figure 1.2: Two tiny graphs describing Markov chains, which are not ergodic.

Next, for ergodic Markov chains, the *mixing time* of a Markov chain M is

$$\tau_M(\varepsilon) := \sup_{x \in \Omega} \min\{t : \|\mathcal{M}^t_{x,\cdot} - \pi\|_{TV} \le \varepsilon\},$$

where $\mathcal{M}$ is the transition matrix of M, $\mathcal{M}^t_{x,\cdot}$ is the row, associated to x of the t-th power of $\mathcal{M}$, $\chi + S$ is the characteristic vector if S, and $\|\cdot\|_{TV}$ is the *variation distance* and defined via:

$$\|\mathcal{M}^t_{x,\cdot} - \pi\|_{TV} := \sup_{S \subset \Omega} |\chi_S \cdot \mathcal{M}^t_{x,\cdot} - \pi|$$

It is the rate of convergence against the unique stationary distribution is what we aim for. Here, a small bound can be used. It allows to use methods like coupling from the past (or at least indicates, it might be feasible to use these methods). So small bounds give us a mean, to sample random elements from Ω. Also good bounds on the mixing time of a Markov chain give insight into the size of the state space Ω. Section 3.2 of [33] covers this reduction, but it is not the scope of this work.

As we see, it is rather hard to find good bounds to the mixing time. In general, we are happy if we are able to decide if a Markov chain M is *rapidly mixing*, which means we can bound $\tau_M(\varepsilon)$ from above by some polynomial in $\log(|\Omega|)$ and $\log(\varepsilon^{-1})$. On the other hand, if we can bound it from below with something growing faster, than such a polynomial, we say M is *not rapidly mixing*. Some examples, where the rapidly mixing property is shown or disproven can be found in Jerrum's book [33]. Note that coupling is a usual method to show that a Markov chain has this property, while the notion of *conductance* is used to disprove it. We follow the same route with this work and use variations of both techniques.

In later chapters we introduce (Markovian) coupling methods, which we use to find such bounds on $\tau_M(\varepsilon)$. So let us stop for now with pointing towards a general introduction to probability: Hägström's book *Finite Markov Chains and Algorithmic Applications* [28].

2 Block Coupling

As announced beforehand, this chapter covers a coupling method, namely *block coupling* and its application to k–heights. These k–heights are related to α-orientations of plane graphs as follows: Reconsider the definition of an α-potential in the previous chapter. An α-potential was an assigned of the graphs faces f to some integer $\phi(f)$, such that

- $|\phi(f) - \phi(f')| \leq 1$ for all $f, f' \in F$, which share an edge e
- $\phi(f) \leq 1$ if f is the only bounded face containing an edge e
- if $f^{l(e)}$ and $f^{r(e)} \in F$ are the faces left and right of e in $\vec{X}_{\min}$, we have $\phi(f^{l(e)}) \leq \phi(f^{r(e)})$

Note, that here the notion of an essential cycle was already replaced by the notion of a face, which suffices for our needs here. The first condition defines some kind of *smoothness* on the mapping. The second one ensures that the mappings values cannot grow arbitrarily, the third notion might look slightly unnatural (although it is not). So to get started, we dropped this third condition and additionally bounded the admissible values of the assignments. The resulting structure is, what you find in the following definition.

2.1 Preliminaries

Definition 5
Let G be a plane graph (a planar graph with a fixed embedding). A *k–height* on G is a mapping

$$h : F(G) \longrightarrow \{0, 1, \dots, k\}$$

such that $|h(f) - h(f')| \leq 1$ for all adjacent faces $f, f' \in F(G)$, i.e. faces of G, which share an edge. The *weight* of a height h is defined as

$$\omega(h) := \sum_{f \in F(G)} h(f).$$

Furthermore, Ω_G (or Ω for short) is the set of all k-heights of G.

Figure 2.1 shows a 2-height on a graph with 13 faces.
Note that for $k = 0$, independent of G, there is only one k-height, which is zero on all faces. For $k = 1$, every $\{0, 1\}$ assignment is valid. So there are $2^{|F|}$ such k-heights and they (as a set) are isomorphic to the set of $\{0, 1\}$ vectors of length $|F|$. For $k = 2$ the game gets interesting, because this is the first instance, where in fact not all possible assignments are valid. A face with value 0 cannot be next to a face with value 2. Luckily, for $k = 2$ or even bigger k, the set of k-heights, has some structure. To get more insight into this structure of the set of all k-heights of a fixed graph, the next lemma is a key ingredient:

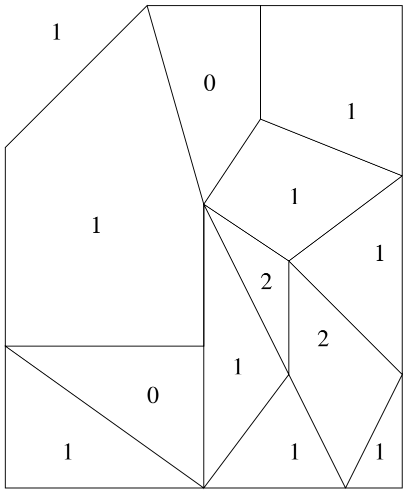

Figure 2.1: A 2-height on a graph with 12 faces.

Lemma 1 *(Ω is a distributive lattice)*
The set of k-heights Ω of a plane graph G together with the face-wise maximum and minimum operations, i.e.

$$\begin{aligned}(h \vee g)(f) &:= \max\{h(f), g(f)\}\\ (h \wedge g)(f) &:= \min\{h(f), g(f)\}\end{aligned}$$

forms a distributive lattice.

Proof. First, $(h \vee g)$ is indeed a k-height. To check this, let f and f' be faces, which share an edge. Without loss of generality, assume $h(f) \geq g(f)$. Now $|(h \vee g)(f) - (h \vee g)(f')| = |\max\{h(f), g(f)\} - \max\{h(f'), g(f')\}| = |h(f) - \max\{h(f'), g(f')\}|$, which is ≤ 1 if $h(f') \geq g(f')$, because h is a k-height.
Assuming $|(h \vee g)(f) - (h \vee g)(f')| > 1$ we know $g(f') > h(f')$, which implies $g(f') \geq h(f)$ and we get $1 > |h(f) - g(f')| = g(f') - h(f) \leq g(f') - g(f) \leq 1$, which contradicts that g is a k-height. With the same arguments, $h \wedge g$ is also a k-height.
Second, the symmetry of $\vee$ and $\wedge$ is inherited from the one of max and min, as well as distributivity, because $\max(a, \min(b, c)) = \min(\max(a, b), \max(a, c))$ for real a, b, and c. □

This directly implies, $h \leq g$ if and only if $h(f) \leq g(f)$ for all faces $f \in F(G)$. Further, the *cover relations* of Ω are exactly pairs of k-heights, which are identical up to one face f, where they differ by 1. These cover relations will be exactly the transitions of the up/down Markov chain M. We can even define a natural metric d for two k-heights h and g:

$$d(h, g) := \sum_{f \in F(G)} |h(f) - g(f)|,$$

which corresponds to the shortest path consisting of cover relations from h to g.

Lemma 2 *(a bound for the maximum distance D in Ω)*
Let D be the maximum distance in Ω with respect to d.

$$D \le k \cdot |F(G)|$$

Proof. Note that between any two k-heights $h, g \in \Omega$, there is a path of cover relations of Ω (or transitions of the up/down Markov chain M), of length

$$d(h,g) = \sum_{f \in F(G)} |h(f) - g(f)| \le \sum_{f \in F(G)} k = k \cdot |F(G)|.$$

This implies $D \le k \cdot |F(G)|$.

□

This Markov chain is somewhat the natural one, operating on the set of k-heights Ω of a fixed plane graph G:

Definition 6 (the Markov chain)
The *up/down Markov chain M* proceeds as follows on a k-height $X \in \Omega$:

```
Input: h ∈ Ω
Output: h′ ∈ Ω
let p ⟵ [0,1] uniformly at random
choose f ⟵ F(G) uniformly at random
choose d ⟵ {+1,−1} uniformly at random
if  p ≤ 1/2 then
|  return h′ = h
else
|  define h′ via h′(f′) := { h(f′) + d   if f′ = f
|                          { h(f′)       else
|  if  h′ ∈ Ω then
|  |  return h′
|  else
|  └  return h′ = h
└
```

Algorithm 1: One step of up/down Markov chain M.

Starting here, we try to analyze the behaviour of M in the remaining pages. It is easy to see, that M is aperiodic, because it is *lazy*, i.e. for every state, with a probability $\ge \frac{1}{2}$ it does not move to another state. Further, it is positive recurrent, because the set of transitions equals the set of cover relations of the distributive lattice Ω, so the transitions of the chain span the whole state space Ω. So M is ergodic and converges against a unique distribution. Further, M is symmetric, i.e. $\Pr(X_{t+1} = g \mid X_t = h) = \Pr(X_{t+1} = h \mid X_t = g)$ for all $h, g \in \Omega$, which implies that $(1, 1, \dots, 1) \in \mathbb{R}^{|\Omega|}$ is an eigenvector with eigenvalue 1 of the transition matrix $\mathcal{M}$ of M. So the unique distribution π the Markov chain M converges to is the uniform distribution.

To bound the mixing time $\tau_M(\varepsilon)$ of M for at least some families of graphs, we make use of *block dynamics* (or *block coupling*), a speed up technique for Markov chains.

For example, Van den Berg and Brouwer applied block dynamics to the monomer-dimer model, i.e. weighted samplings of graphs, see [7] for details. The idea behind block coupling is, to introduce blocks, which induce a decomposition of the state space Ω. Based on this, a block Markov chain is defined, which chooses one of the blocks and chooses from the set of all elements of the state space, which are identical to the current one outside of the block and pick one of them, i.e. the block's content is updated completely in one step, while the outside remains the same.
Assuming the blocks and the block Markov chain are *nice*, comparison theorems can be used to connect the block Markov chain's behaviour to the behaviour of the original chain. In our case, this looks as follows:

Definition 7 (block)
A *block* B is a set of faces of G. Further, ∂B is the *border* of B, i.e. all $f \in F(G) \setminus B$, which share an edge with one face of B.

So from now on, let $\mathcal{B}$ be a finite set of finite Blocks B living on a fixed and plane graph G, such that every face of G is contained in at least one block.
Now, Ω_B is the set of k-heights on $B \in \mathcal{B}$ and let $\Omega_{\partial B}$ be the set of k-heights on ∂B. They form the set of all valid *border conditions* for interior fillings of ∂B, i.e. B. Further, $\Omega_{B|h} \subset \Omega_B$ is the set of *admissible allocations* of B with respect to h, i.e. k-heights on B, which are admissible with respect to the values of h on ∂B. Let us encounter this with an example:

Example 1
Consider the block in Figure 2.2. It consists of three (disconnected) faces a, b, and c and has a border, which consists of nine faces.

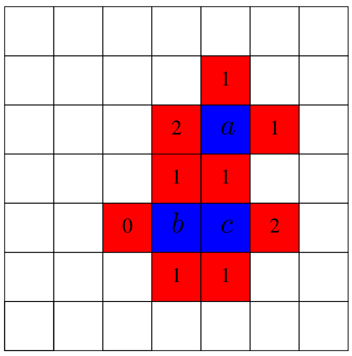

Figure 2.2: The block B consisting of 3 blue faces a, b, and c together with the values of h on the 9 red faces of ∂B in a subsection of the rectangular lattice.

The set Ω_B contains 21 elements (of which not all conform to the border conditions, imposed by h). Figure 2.3 shows them as a distributive lattice in the notion, where h is represented by $h(a)h(b)h(c)$.
Now let h be a two-height on ∂B such that $h(f)$ equals the values, indicated in Figure 2.2.
Now there are six admissible allocations on B respecting h. These six are 101, 201, 111, 211, 112, and 212, so $\Omega_{B|h}$ is the set of these six 2-heights on B.

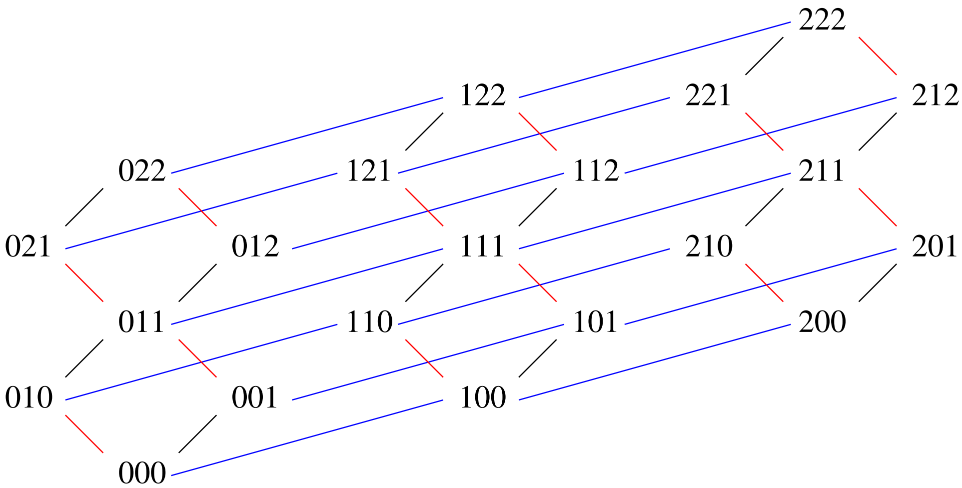

Figure 2.3: The distributive lattice Ω_B of twenty-one 2-heights. Changes of a are blue edges, changes of b red and changes of c black.

Now, after establishing the notions of blocks and admissible allocations, the block Markov chain can be defined:

Definition 8 (the block Markov chain)
The *block Markov chain* $M_{\mathcal{B}}$ proceeds as follows on a k-height h:

Input: $h \in \Omega$
Output: $h' \in \Omega$
Let $p \longleftarrow [0,1]$ uniformly at random
if $p \leq \frac{1}{2}$ **then**
  return $h' = h$
else
  choose $B \longleftarrow \mathcal{B}$ uniformly at random
  choose $g \longleftarrow \Omega_{B|h}$ uniformly at random
  define h' via $h'(f) := \begin{cases} g(f) & \text{if } f \in B \\ h(f) & \text{else} \end{cases}$
  return h'

Algorithm 2: One step of block Markov chain $M_{\mathcal{B}}$.

First, note that the output is indeed a k-height, because every element $g \in \Omega_{B|h}$ is admissible with the border restrictions, imposed by h, i.e. $|h(f) - g(f')| \leq 1$ for all faces $f \in \partial B$ and $f' \in B$, which share an edge.
Second, for every face $f \in F(G)$ we have a block $B \in \mathcal{B}$, so by choosing this block and the right transition to replace the block, every transitions of the up/down Markov chain M is also a transition of $M_{\mathcal{B}}$. So $M_{\mathcal{B}}$ inherits its reducible and positive recurrence from M. Besides, $M_{\mathcal{B}}$ is also lazy, which implies, it is aperiodic. Finally $M_{\mathcal{B}}$ is symmetric. So together, the block Markov chain $M_{\mathcal{B}}$ is ergodic and the unique distribution the chain converges to is the uniform one on Ω. To conclude, $M_{\mathcal{B}}$ is a well behaving Markov chain to encounter the behaviour of M. Knowing this, we can formulate and prove our main

theorem in the next section.

2.2 The main theorem

This section establishes the main theorem, which will allow us to prove, that the up/-down Markov chain M is rapidly mixing the set of 2-heights for some families of graphs in later sections. This theorem is based on some parameters of the set of blocks $\mathcal{B}$, which the block Markov chain uses. So in later sections, we only encounter these properties of $\mathcal{B}$ and apply our theorem to find a bound for the mixing time of the up/down Markov chain for these graphs.
Two notions, needed to formulate the theorem are $C_{B,\delta}$ and $E_{B,\delta}$. The first, $C_{B,\delta}$, is the set of cover relations of $\Omega_{\partial B}$, which differ in δ, i.e.

$$C_{B,\delta} := \{((l,u) \in \Omega_{\partial B}^2 \mid u(\delta) = l(\delta) + 1 \text{ and } u(f) = l(f) \text{ for all } f \in \partial B \setminus \{\delta\}\}$$

and the second, $E_{B,\delta}$, is formally defined as

$$E_{B,\delta} := \max_{(l,u)\in C_{B,\delta}} \{\mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+))\}.$$

Here, $l^d{+}_B$ is a random variable, uniformly distributed on $\Omega_{B|l}$ and u_B^+ is a random variable, uniformly distributed on $\Omega_{B|u}$. Further, $\omega(u_B^+) = \sum_{f\in B} u_B^+(f)$ is the weight of u_B^+. Later in the proof we will see, that $E_{B,\delta}$ equals the maximum expected height difference in $C_{B,\delta}$, i.e. we will prove $E_{B,\delta} = \max_{(l,u)\in C_{B,\delta}}\{\mathbb{E}(\sum_{f\in B}|u^+(f) - l^+(f)|)\}$. This obfuscation of the definition is more than paid of, because $E_{B,\delta}$ can be much more easily computed. We can compute $\mathbb{E}(\omega(h))$ for each h independently and then compare all elements of $C_{B,\delta}$ according to the precomputed values. If the equation would not hold, we would have to compute $\mathbb{E}(\sum_{f\in B}|u^+(f) - l^+(f)|)$ for all elements of $C_{B,\delta}$. This computation is much slower than the computation of all $\mathbb{E}(\omega(h))$, so breaking down the number of such computations from $|C_{B,\delta}|$ to $|\Omega_B|$ can be quite substantial to actually be able to apply the theorem. This possibility to speed up the computations required for the application of the theorem is an integral feature of it. Finally, the theorem is, as follows:

Theorem 2
Let $\mathcal{B}$ be a finite set of blocks, such that for every face $f \in F(G)$ there is at least one block $B \in \mathcal{B}$ with $f \in B$ and let $\beta \in \mathbb{R}$, such that

$$1 > \beta \geq 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B\in\mathcal{B}|\delta\in\partial B} E_{B,\delta} \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right)$$

for all faces $\delta \in F(G)$. We have

$$\tau_M(\varepsilon) \leq c_{\mathcal{B},k} \cdot \frac{(\log(\frac{1}{\varepsilon}\cdot|F(G)|) + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k\cdot|F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})},$$

where $m := \max_{f\in F(G)} |\{B \in \mathcal{B} \mid f \in B\}|$, $b := \max_{B\in\mathcal{B}}\{|B|\}$, and $c_{\mathcal{B},k} := \frac{16\cdot b\cdot m\cdot k\cdot(k+1)^b}{1-\beta}$.

The proof of this theorem needs some machinery, which we introduce before jumping into the details. Still, the proof's outline is, as follows: In the end, we want to apply some path coupling, so we have to get our hands on the expected change of distance if the coupling is applied to neighbouring pairs of k-heights. The theorem relies only on the expected weight of all admissible allocations for a fixed set of border conditions. To close this gap, we construct a monotone coupling with a discrete version of Strassen's theorem on stochastic dominance. To do so, the Ahlswede-Daykin 4-Functions theorem is applied to show, that two distributions are indeed dominating each other stochastically. Afterwards, some Erdös-Magic, relying on the couplings monotonicity, reveals, that the path coupling theorems of Dyer and Greenhill can indeed be applied. This leads to a bound for the mixing time of the block Markov chain $M_{\mathcal{B}}$. Finally a version of the comparison algorithm of Diaconis and Saloff-Coste yields a bound to the mixing time of the up/down Markov chain M. So before we prepare the stage, let us consider the border cases of the Theorem:

Remark 1
Assume, all blocks consist of one face. Therefore, for every face $f \in F$ there is block $B_f = \{f\} \in \mathcal{B}$. Now, the condition to β implies, that if the theorem was applicable, i.e. there is a such a β, we have

$$\sum_{B \in \mathcal{B} | \delta \in \partial B} E_{B,\delta} \quad < \#\{B \in \mathcal{B} \mid \delta \in B\} = 1.$$

This implies, that the expected change in distance of the up/down Markov chain is negative, easily yielding rapidly mixing of the Markov chain, because Theorem 4 can be applied. This theorem, which we will see in a few pages is one of the tools, needed to prove Theorem 2. Therefore a direct application of it would be quite some shortcut and most likely even improve the resulting bound to the mixing time.

On the other hand, assume $\mathcal{B}$ contains only of a block $B = F$, i.e. we group all faces of G in one block together. Therefore, the condition to β is trivially applicable (because δ can never be on the border of this block) and we can get some $1 > \beta > \frac{1}{2}$. On the other hand, $c_{\mathcal{B},k}$ is exponential in $|F|$. It contains some $(k+1)^{|F|} \geq |\Omega|$, while the denominator is $\geq \frac{1}{2}$. Therefore the bound is linear in something, which is $> |\Omega|$, and never yields some rapidly mixing.

To conclude, the art in the application of Theorem 2 is the choice of the right family of blocks of which one can prove the existence of some β which suffices to the condition and whose block sizes are reasonable small (i.e. constant or logarithmic in terms of $|F|$).

Now, we start to collect all tools to prove the theorem. First comes the theorem below, which helps in the end to create some monotone coupling:

Theorem 3 *(Ahlswede Daykin 4-Functions Theorem; [2])*
Let D be a finite distributive lattice and $f_1, f_2, f_3, f_4 : D \to \mathbb{R}_{\geq 0}$, such that

$$f_1(a) \cdot f_2(b) \leq f_3(a \vee b) \cdot f_4(a \wedge b) \text{ for all } a, b \in D.$$

This implies

$$f_1(A) \cdot f_2(B) \leq f_3(A \vee B) \cdot f_4(A \wedge B) \text{ for all } A, B \subseteq D,$$

where $f_i(A) = \sum_{a \in A} f_i(a)$ *and* $A \vee B = \{a \vee b \mid a \in A, b \in B\}$ *and* $A \wedge B = \{a \wedge b \mid a \in A, b \in B\}$.

To make use of this, we need the notion of stochastic dominance:

Definition 9 (stochastic dominance)
Let d_1 and d_2 be distributions on a finite and partially ordered set Ω. d_1 *is stochastically dominated by* d_2, in symbols $d_1 \leq_{\text{stoch}} d_2$, if and only if

$$\sum_{a \in F} d_1(a) \leq \sum_{a \in F} d_2(a)$$

for all upsets $F \subseteq \Omega$.

Now, using this definition of stochastic dominance, we will show and prove a version of Strassen's Theorem on stochastic dominance:

Proposition 1 *(A discrete version of Strassen's theorem on stochastic dominance)*
Let d_1, d_2 *be random distributions on the finite partially ordered set* Ω. *Now* $d_1 \leq_{stoch} d_2$*implies the existence of such a distribution* $q(x,y)$ *on* $\Omega \times \Omega$ *with*

$$q(x,y) > 0 \textit{ implies } x \leq y$$

and

$$\sum_{y \in \Omega} q(x,y) = d_1(x) \textit{ for all } x \textit{ and } \sum_{x \in \Omega} q(x,y) = d_2(y) \textit{ for all } y.$$

This is a special case of Strassen's theorem on stochastic dominance, [47]. We will present a rather short proof, which relies on the MaxFlow-MinCut - Theorem. Another proof for a discrete version of Strassen's theorem is contained in [4]. This one is based on Farkas' Lemma. The formulation and a proof sketch of this theorem can also be found in Problem set A of Steve Lalley's Summer Course [37], so he –at least– is aware of this version of Strassen's theorem.

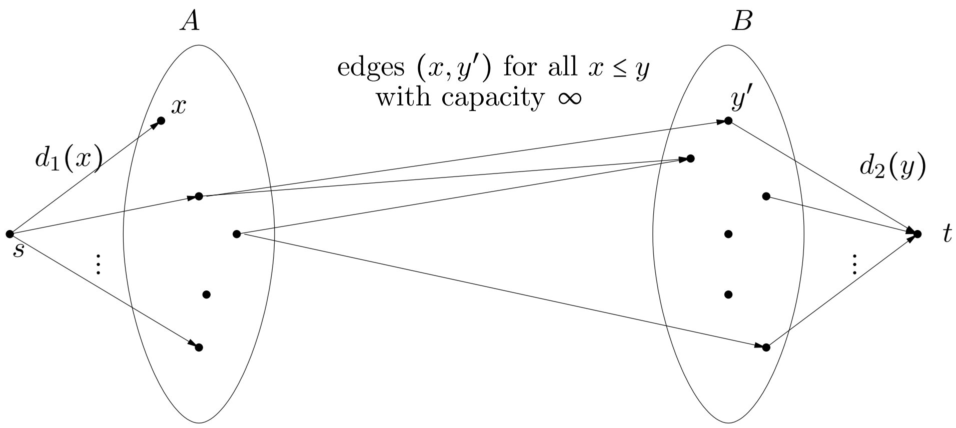

Figure 2.4: The network-flow N.

Proof. Consider the network N in Figure 2.4. It consists of a source s, a sink t and two blocks of vertices $A := \{x \mid x \in \Omega\}$ and $B := \{y' \mid y \in \Omega\}$. Further, s is connected to every $x \in A$ with an edge of capacity $d_1(x)$ and every vertex $y' \in B$ is connected to t

with an edge of capacity $d_2(y)$. Finally, there is an edge (x, y') with $x \in A$ and $y' \in B$ of capacity ∞, if $x \leq y$.

Let us assume, there is a network flow from s to t of value 1. This implies, the flow uses the complete capacity of all edges, to which s is adjacent to, because the sum of their weights is $\sum_{x \in A} d_1(x) = 1$, because d_1 is a distribution. Obviously, the same holds for all edges, which end in t.
Let $q(x, y)$ be the value of the flow on the edge (x, y'). Obviously, $q(x, y) = 0$ if $x \nleq y$, because F does not contain the edge (x, y'). Further, $\sum_{y \in \Omega} q(x, y) = d_1(x)$ for all $x \in \Omega$, because the incoming flow $d_1(x)$ has to match the outgoing flow $\sum_y q(x, y)$ at the vertex x. By symmetry, we also get $\sum_x q(x, y) = d_2(y)$ for all $y \in \Omega$. So if the assumption was correct, the proposition is true.

To prove the assumption, first note that any (s, t)-cut C with capacity $< \infty$ cannot contain any edge between A and B. Still, for every edge (x, y') between A and B the (s, t)-cut C has to contain either (s, x) or (y', t), otherwise it does not separate s from t. Due to $x \leq x'$, the edge (x, x') exists in N, so C has to contain (s, x) or (x', t). So every (s, t)-cut with finite capacity induces a vertex cover of the bipartite graph N', which consists of N without s, t and the edges, which contain s and t.

Next, consider special vertex covers, which are induced by upsets $F \subseteq \Omega$. So let F be an upset of Ω. Now the set $C := \{x \mid x \notin F\} \cup \{y' \mid y \in F\}$ is a vertex-cover of N', which induces an (s, t)-cut of N: For all $x \in F$ and $y \in \Omega \setminus F$ we have $x \nleq y$, so (x, y') is no edge of N'. This implies, that all edges (x, y') of N' contain at least one vertex of C, so the set of edges $\{(s, x) \mid x \notin F\} \cup \{(y', t) \mid y \in F\}$ indeed forms an (s, t)-cut. Further, the capacity of this cut is $1 - p(F) + q(F)$, and we even now that this is ≥ 1, because $p \leq_{\text{stoch}} q$ is a requirement and F is an upset, so $p(F) \leq q(F)$.

Now, let C be an arbitrary (s, t)-cut with finite capacity. Let $C_L \subseteq A$ be the set of vertices in A being incident to an edge of the cut C. Formally, we have $C_L := \{x \in \Omega \mid (s, x) \in C\}$. In the same manner and $C_R := \{y' \mid y \in \Omega \text{ and } (y', t) \in C\}$ is the set of vertices in B which touch an edge of C. We want to show the existence of an upset U, which induces a vertex cover, such that its cut has –at least– no bigger capacity. To do so, consider $L = A \setminus C_L$ and $R := B \setminus C_R$. Further, let $U = \{y \in \Omega \mid \exists x \in L : x \leq y\}$ be the upset spanned by L and $U' := \{y' \mid y \in \Omega \text{ and } \exists x \in L : x \leq y\}$ its image in B, see Figure 2.5.
Now, we claim $U' \subseteq C_R$. To argue for this, consider an element $y' \in U' \setminus C_R$. By definition of U' there is some $x \in L$ with $x < y$. So there is an edge $(x, y') \in N$ and neither x nor y' touch an edge of C. This implies we have the path $s, (s, x), x, (x, y'), y', (y', t)$ in $N \setminus C$, so C is not a (s, t)-cut. So the claim is true an we have $U' \subseteq C_R$.

Now, the capacity of the cut induced by U is not bigger than the capacity of C:

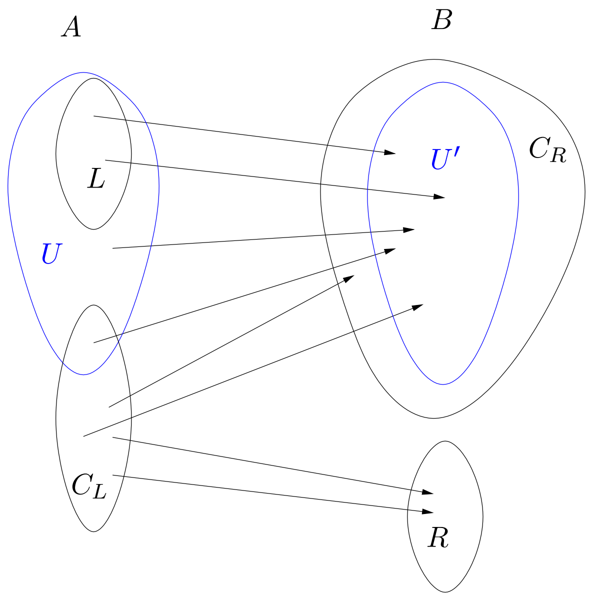

Figure 2.5: The subdivision of A and B into C_L, L, U, C_R, R and U'. There are no edges from U to R.

$$\begin{aligned}
\mathrm{cap}(C) &= \sum_{x\in C_L} d_1(x) + \sum_{y\in C_R} d_2(y) \\
&= \sum_{x\in C_L\setminus U} d_1(x) + \sum_{x\in C_L\cap U} d_1(x) + \sum_{y\in C_R\setminus U} d_2(y) + \sum_{y\in C_R\cap U} d_2(y) \\
&\geq \sum_{x\in C_L\setminus U} d_1(x) + \sum_{y\in C_R\cap U} d_2(y) \\
&= \sum_{x\notin U} d_1(x) + \sum_{y\in U} d_2(y) \\
&= \mathrm{cap}(U),
\end{aligned}$$

This implies, that for every cut C, there is a cut U, induced by an upset, which does not have a bigger capacity. Besides, every cut, induced by an upset, has a capacity ≥ 1. Finally, there is a (s,t)-cut with capacity 1 (for example the set of edges, adjacent to t), so we conclude that 1 is the capacity of the minimum cut. Therefore, by the MaxFlow=MinCut-Theorem we know, that there is a maximum flow with value 1 and the assumption is true. This finishes this proof. □

Next, we have to make sure, that we indeed have the case of two distributions, stochastically dominating each other if we have border restrictions, which are comparable:

Lemma 3 *(stochastic dominance)*
Let $l, u \in \Omega_{\partial B}$ with $l \leq u$ be border conditions, i.e. k-heights on ∂B.

Further, let d_l and d_u be the corresponding distributions on Ω_B, i.e.

$$d_u(h) := \begin{cases} \frac{1}{|\Omega_{B|u}|} & \text{if } h \in \Omega_{B|u} \\ 0 & \text{else} \end{cases} \quad \text{and}$$

$$d_l(h) := \begin{cases} \frac{1}{|\Omega_{B|l}|} & \text{if } h \in \Omega_{B|l} \\ 0 & \text{else} \end{cases}$$

for all $h \in \Omega_B$. This implies

$$d_l \leq_{stoch} d_u.$$

To prove this lemma, we need another one, which we formulate and prove for starters:

Lemma 4
Let $D \subseteq \Omega_B$ be the smallest distributive lattice, which contains $\Omega_{B|l} \cup \Omega_{B|u}$. Now, $\Omega_{B|l}$ is a down set in D and $\Omega_{B|u}$ forms an upset.

Proof. Let $h \in \Omega_{B|l}$ and $g \in D$ such that $g \leq h$. We have to show $g \in \Omega_{B|l}$. Assume $g \notin \Omega_{B|l}$. We have $g \in D \subset \Omega_B$, so g is a valid k-height, but does not comply with the border conditions, imposed by l. This means there is a face $f \in B$ sharing an edge with a face $f' \in \partial B$, such that $|g(f) - l(f')| > 1$. Even more, we know $g(f) < l(f') - 1$, because $g < h \in \Omega_{B|l}$ implies $g(f) \leq h(f) \in \{l(f') - 1, l(f'), l(f') + 1\}$.
On the other hand, we know $u \geq l$, which yields $u(f') \geq l(f')$. This means, for all $h' \in \Omega_{B|l} \cup \Omega_{B|u}$ we have $h'(f) \geq l(f') - 1$, implying $(\min D)(f) \geq l(f') - 1$, because $(\min D)(f') = \min\{h'(f') \mid h' \in \Omega_{B|l} \cup \Omega_{B|u}\}$. So $g(f) < (\min D)(f)$, which contradicts $g \in D$!
To show that $\Omega_{B|u}$ is an upset in D, the arguments above suffice, due to the symmetry of all conditions. □

Now it is time to prove the lemma on stochastic dominance stated above:

Proof. (of Lemma 3)
Due to $d_u(h) = d_l(h) = 0$ for all $h \in \Omega_B \setminus (\Omega_{B|u} \cup \Omega_{B|l})$ we can restrict ourselves to the distributive lattice D, spanned by $\Omega_{B|u} \cup \Omega_{B|l}$. In this lattice, $\Omega_{B|u}$ is an upset and $\Omega_{B|l}$ forms a down set, according to Lemma 4.
To show $d_l \leq_{\text{stoch}} d_u$, let F be an upset of D. We have to show

$$\begin{aligned} 0 &\leq \sum_{a \in F} u(a) - \sum_{a \in F} l(a) \\ &= \frac{|F \cap \Omega_{B|u}|}{|\Omega_{B|u}|} - \frac{|F \cap \Omega_{B|l}|}{|\Omega_{B|l}|}, \end{aligned}$$

which is equivalent to

$$(*) \qquad 0 \leq |F \cap \Omega_{B|u}| \cdot |\Omega_{B|l}| - |F \cap \Omega_{B|l}| \cdot |\Omega_{B|u}|.$$

To prove this equation, we apply Theorem 3, the Ahlswede-Daykin 4-functions theorem. So we need four functions. These are defined by

$$\begin{aligned} f_1(h) &:= \chi_{F\cap\Omega_{B|l}}(h) \\ &= \begin{cases} 1 & \text{if } h \in F \cap \Omega_{B|l} \\ 0 & \text{else} \end{cases}, \\ f_2(h) &:= \chi_{\Omega_{B|u}}(h), \\ f_3(h) &:= \chi_{F\cap\Omega_{B|u}}(h), \text{ and} \\ f_4(h) &:= \chi_{\Omega_{B|l}}(h) \end{aligned}$$

for all $h \in D$. Now, for $h, g \in D$ with $f_1(h) \cdot f_2(g) = 1$ we know $h \in F \cap \Omega_{B|l}$ and $g \in \Omega_{B|u}$. This implies $h \vee g \in \Omega_{B|u}$, because $h \vee g \geq g$ and $\Omega_{B|u}$ forms an upset. It also implies $h \vee g \in F$, because F is again an upset and $h \in F$.
Together it yields $h \vee g \in F \cap \Omega_{B|u}$, i.e. $f_3(h \vee g) = 1$. On the other hand $h \wedge g \leq h \in \Omega_{B|l}$ and $\Omega_{B|l}$ is a down set, so we have $h \wedge g \in \Omega_{B|l}$ and $f_4(h \wedge g) = 1$. Further, $f_1(h) \cdot f_2(g) \in \{0, 1\}$ for all $h, g \in D$. This implies:

$$f_1(h) \cdot f_2(g) \leq f_3(h \vee g) \cdot f_4(h \wedge g) \text{ for all } h, g \in D$$

This means, we can apply the 4-functions theorem. Doing so with $A = B = D$ yields:

$$\begin{aligned} 0 &\leq f_3(D \vee D) \cdot f_4(D \wedge D) - f_1(D) \cdot f_2(D) \\ &= f_3(D) \cdot f_4(D) - f_1(D) \cdot f_2(D) \\ &= \chi_{F\cap\Omega_{B|u}}(D) \cdot \chi_{\Omega_{B|l}}(D) - \chi_{F\cap\Omega_{B|l}}(D) \cdot \chi_{\Omega_{B|u}}(D) \\ &= |F \cap \Omega_{B|u}| \cdot |\Omega_{B|l}| - |F \cap \Omega_{B|l}| \cdot |\Omega_{B|u}|, \end{aligned}$$

which is equation (*). So (*) holds, which implies that u stochastically dominates l. □

With this lemma, we are able to establish our coupling. It tells us, that we can apply Proposition 1 to two border conditions u and l, such that $u \geq l$. This yields a distribution q, which will allow us to successfully couple two instances of our block Markov chain $M_{\mathcal{B}}$.
Before doing so, we first want to state three more theorems, which we need later to bound the mixing time of the block Markov chain, via the coupling we can construct by the application of Strassen's theorem.

Theorem 4 *(Dyer and Greenhill;first part of Theorem 2.1, [17])*
Let (X_t, Y_t) be a coupling for the Markov chain M and let d be any integer valued metric defined on $\Omega \times \Omega$. Suppose that there exists $\beta < 1$ such that

$$\mathbb{E}[d(X_{t+1}, Y_{t+1})] \leq \beta \cdot d(X_t, Y_t)$$

for all t. Let D be the maximum value that d achieves on $\Omega \times \Omega$. The mixing time $\tau_M(\varepsilon)$ of M satisfies

$$\tau_M(\varepsilon) \leq \frac{\log(D \cdot \frac{1}{\varepsilon})}{(1 - \beta)}.$$

Theorem 5 *(Dyer and Greenhill;Theorem 2.2, [17])*
Let d be an integer-valued metric defined on $\Omega \times \Omega$ which takes values in $\{0, ..., D\}$. Let S be a subset of $\Omega \times \Omega$ such that for all $(X_t, Y_t) \in \Omega \times \Omega$ there exists a path $X_t = Z_0, Z_1, \dots, Z_r = Y_t$ between X_t and Y_t such that $(Z_l, Z_{l+1}) \in S$ for $0 \leq l < r$ and

$$\sum_{l=0}^{r-1} d(Z_l, Z_{l+1}) = d(X_t, Y_t).$$

Define a coupling $(X, Y) \mapsto (X', Y')$ of the Markov chain M on all pairs $(X, Y) \in S$. Apply this coupling along the given sequence from X_t to Y_t to obtain the new sequence $Z_{0'}, Z_{1'}, \dots, Z_{r'}$. Therefore (X_t, Y_t) is also a coupling for M, where $X_{t+1} = Z_{0'}$ and $Y_{t+1} = Z_{r'}$. Moreover, if there exists $\beta \leq 1$ such that $\mathbb{E}[d(X', Y')] \leq \beta \cdot d(X, Y)$ or all $(X, Y) \in S$ we have

$$\mathbb{E}[d(X_{t+1}, Y_{t+1})] \leq \beta \cdot d(X_t, Y_t).$$

Theorem 6 *(Randall and Tetali, Proposition 4 of [42])*
Let M and $\widetilde{M}$ be two reversible Markov chains on the same state space Ω with the same stationary distributions π, and let $\pi_\star := \min_{x \in \Omega} \pi(x)$. Let $E(M)$ be the set of transitions of M and $E(\widetilde{M})$ be the set of transitions of $\widetilde{M}$. For each pair $(u, v) \in E(\widetilde{M})$ define a path γ_{uv}, which is a sequence $u = u_0, u_1, \dots, u_k = v$ of transitions of M, i.e. $(u_i, u_{i+1}) \in E(M)$ for all i. For $(x, y) \in E(M)$ let

$$\Gamma(x, y) := \{(u, v) \in E(\widetilde{M}) \mid (x, y) \in \gamma_{uv}\}.$$

Further let

$$A := \max_{(x,y) \in E(M)} \left\{ \frac{1}{\pi(x) \cdot M(x,y)} \cdot \sum_{(u,v) \in \Gamma(x,y)} |\gamma_{uv}| \cdot \pi(u) \cdot \widetilde{M}(u, v) \right\},$$

where $M(x, y) := \Pr_M(X_{t+1} = y \mid X_t = x)$, i.e. the transition probability from x to y in one step / the entry associated to x and y in the transition matrix $\mathcal{M}$ of M. If the second largest eigenvalue λ_1 of $\widetilde{M}$ complies with $\lambda_1 \geq \frac{1}{2}$, we have for every $0 < \varepsilon < 1$ the mixing time $\tau_M(\varepsilon)$ is bounded by:

$$\tau_M(\varepsilon) \leq \frac{4 \log(\frac{1}{\varepsilon \cdot \pi_\star})}{\log(\frac{1}{2\varepsilon})} \cdot A \cdot \tau_{\widetilde{M}}(\varepsilon)$$

As announced, after these three theorems we can finally define our coupling P of the block Markov chain $M_{\mathcal{B}}$:

Definition 10 (the coupling)
The coupling P on $\Omega \times \Omega$ transforms $(h, g) \in \Omega \times \Omega$ as follows:

Input: $(h, g) \in \Omega \times \Omega$
Output: $(h', g') \in \Omega \times \Omega$
choose $p \longleftarrow [0, 1]$ uniformly at random
if $p \leq \frac{1}{2}$ **then**
  return $(h', g') \longleftarrow (h, g)$
else
  choose a block $B \longleftarrow \mathcal{B}$ uniformly at random
  if $h_{|\partial B}$ *equals* $g_{|\partial B}$ **then**
    `// we have` $\Omega_{B|\partial h} = \Omega_{B|\partial g}$
    choose $(x, y) \in \Omega_{B|\partial h} \times \Omega_{B|\partial g}$ with $x = y$ uniformly at random
  else
    if $\{h_{|\partial B}, g_{|\partial B}\}$ *is a cover relation of* $\Omega_{\partial B}$ **then**
      the uniform distributions d_g, d_h of the admissible allocations on $L = \text{span}_{\text{dist.lattice}}\{\Omega_{B|h} \cup \Omega_{B|g}\}$ stochastically dominate each other according to Lemma 3. This means Proposition 1 can be applied, so
      choose $(x, y) \in \Omega_{B|h} \times \Omega_{B|g}$ according to the distribution $q(h_{|\partial B}, g_{|\partial B})$ of Proposition 1.
    else
      determine (x, y) according to Theorem 5
  define h' via $h'(f) := \begin{cases} x(f) & \text{if } f \in B \\ h(f) & \text{else} \end{cases}$
  define g' via $g'(f) := \begin{cases} y(f) & \text{if } f \in B \\ g(f) & \text{else} \end{cases}$
  return (h', g')

Algorithm 3: One step of the coupling P of the block Markov chain $M_{\mathcal{B}}$.
Here, $h_{|\partial B}$ is the restriction of h to the faces of ∂B, so $\{h_{|\partial B}, g_{|\partial B}\}$ is a cover relation if and only if there is a face $\delta \in \partial B$, such that $h(f) = g(f)$ for all $f \in \partial B \setminus \{\delta\}$ and $h(\delta) \in \{g(\delta) - 1, g(\delta) + 1\}$.

Example 2 (an example of a coupling step)
To exemplify, how P works, let us consider an example. Let B be the block of three faces in Figure 2.6 with a border of seven faces.

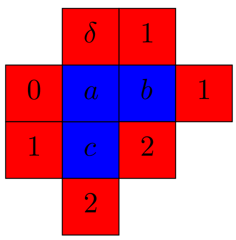

Figure 2.6: The block B of three blue faces with seven red faces on the border.

Let l and u be two border conditions for B, i.e. 2-heights on ∂B. Their values on the

faces is given in Figure 2.6. Further, let $l(\delta) = 1$ and $u(\delta) = 2$.
Now, write a k-height h as $(h(a), h(b), h(c))$. The lattice L, spanned by the sets of admissible allocations $\Omega_{B|l}$ and $\Omega_{B|u}$ is $L = ((0,1,1,), (1,1,1), (1,2,1), (1,1,2), (1,2,2))$, as shown in Figure 2.7.

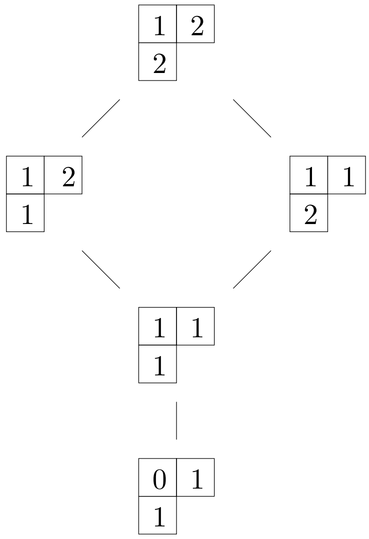

Figure 2.7: The distributive lattice of k-heights on B.

Since $l \leq u$, we make use of Proposition 1 in the coupling P. Here, the distributions are $d_l = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$ and $d_u = (0, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ (indexed by the order in L). It is easy to check, that indeed $d_l \leq_{\text{stoch}} d_u$, because L has only five upsets, which we have to consider. Setting up the network flow N for this lattice L and and finding a flow of value 1 yields a (not unique) distribution

$$q := \begin{pmatrix} 0 & \frac{1}{20} & \frac{1}{20} & \frac{1}{20} & \frac{1}{20} \\ 0 & \frac{1}{5} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{pmatrix}.$$

It means, that (u, l) is updated to (u^+, l^+) on B with probability q_{u^+, l^+}. Note that the blue zeros are imposed by the condition $x \nleq y \Rightarrow q(x, y) = 0$.

This coupling leads to $\mathbb{E}(d(l^+, u^+)) = \frac{1}{20} \cdot 1 + \frac{1}{20} \cdot 2 + \frac{1}{20} \cdot 2 + \frac{1}{20} \cdot 3 = \frac{2}{5} < 1$. If for all other blocks $B' \in \mathcal{B}$ and and all other pairs of border conditions this expectation is ≤ 1, the main theorem would show that the up/down Markov chain is indeed rapidly mixing.

Still it is not clear, how the conditions stated in Theorem 2 suffice to apply Theorem 5, which requires, that the expected distance after one coupling step, applied to a cover relation is < 1. To establish this link, consider the following lemma:

Lemma 5 *(expected distance after one coupling step applied to a cover relation)*
Let $\delta \in F(G)$ be a face and $l < u$ be two k-heights, such that $l(\delta) + 1 = u(\delta)$ and $l(f) = u(f)$ for all $f \in \partial F(G) \smallsetminus \{\delta\}$, i.e. (l, u) is a cover relation of Ω. Further, let

(l^+, u^+) be a pair of random variables, distributed according to the coupling P applied to (l,u) and

$$\beta \geq 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B\in\mathcal{B}|\delta\in\partial B} E_{B,\delta} \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right)$$

for all faces $\delta \in F(G)$, as postulated in Theorem 2. This implies

$$\mathbb{E}(d(u^+, l^+)) \leq \beta.$$

Proof. First, remember the definition of $E_{B,\delta}$:

$$E_{B,\delta} := \max_{(l,u)\in C_{B,\delta}} \left(\mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+))\right),$$

where $C_{B,\delta}$ is the set of cover relations of $\Omega_{\partial B}$, which differ in δ. This implies $E_{B,\delta} \geq \mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+))$ for our fixed cover relation (l,u), so we have

$$\beta \geq 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B\in\mathcal{B}|\delta\in\partial B} \mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+)) \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right)$$

Next, let us assume that the coupling P chooses a block B, such that $\delta \in \partial B$. So let $(l_B^+, u_B^+) \in \Omega_{B|l} \times \Omega_{B|u}$ be a pair of random variables, distributed according to q of Proposition 1.
Now $l < u$ ensures $(l_B^+) \leq (u_B^+)$, due to the construction of the coupled distribution q. So we have $l_B^+(f) \leq u_B^+(f)$ for all faces $f \in B$ and therefore

$$|u_B^+(f) - l_B^+(f)| = u_B^+(f) - l_B^+(f) \text{ for all } f \in B.$$

This implies:

$$\begin{aligned}
\mathbb{E}(d(u_B^+, l_B^+)) &= \mathbb{E}(\sum_{f\in F(G)} |u_B^+(f) - l_B^+(f)| \\
&= \mathbb{E}(\sum_{f\in B} u_B^+(f) - l_B^+(f)) + 1 \\
&= \mathbb{E}(\sum_{f\in B} u_B^+(f)) - \mathbb{E}(\sum_{f\in B} l_B^+(f)) + 1 \\
&= \mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+)) + d(u,l)
\end{aligned}$$

Note, that the coupling P does not change the values of the pair of k-heights outside of the block B (or on its border) at all, so $\mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+))$, which is only estimated inside of B, captures the complete *change* of the distance. Further, that u_B^+ is distributed uniformly in $\Omega_{B|u}$, because P is a coupling of the up/down Markov chain M, which chose it uniformly at random from $\Omega_{B|u}$. Therefore, the $\mathbb{E}(u_B^+)$, which occurs here is the same as in the definition of $E_{B,\delta}$! Obviously, the same holds for l_B^+ and $\mathbb{E}(l_B^+)$, so we move here from a pair of coupled random variables to two independent random variables, both uniformly distributed on their corresponding sets of k-heights! This crucial step enables us, to find bounds for $E_{B,\delta}$ later for at least some families of blocks.

To prove the claim, note, that whenever P chooses a block B, such that $\delta \in B$, the coupling returns a pair of equal k-heights, so the distance is decreased by 1. On the other hand, of δ is neither in B nor in ∂B, the distance is not changed by the coupling. Finally, if and only if $\delta \in \partial B$, the distance might be increased by $\mathbb{E}(u_B^+) - \mathbb{E}(l_B^+)$. These considerations lead to:

$$\begin{aligned}
\mathbb{E}(d(u^+, l^+)) &= 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B \in \mathcal{B} | \delta \in \partial B} \mathbb{E}(d(u_B^+, l_B^+)) \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right) \\
&= 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B \in \mathcal{B} | \delta \in \partial B} \mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+)) \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right) \\
&\leq 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B \in \mathcal{B} | \delta \in \partial B} E_{B,\delta} \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right) \\
&\leq \beta
\end{aligned}$$

So the claim is also true and the expected distance of a cover relation of Ω after one step of the coupling P is bounded by $\beta < 1$.

□

Now, we collected all tools which are needed to finally prove the main theorem.

Proof. (of Theorem 2)
Let (l, u) be a cover relation of Ω. So u and l are the same except on one face $\delta \in F(G)$, where $l(\delta) + 1 = u(\delta)$. Let (u^+, l^+) be a pair of random variables, resulting in the application of one step of the coupling P to (u, l). Next, Lemma 5 tells us, that the expected distance between u^+ and l^+ is at most $\beta < 1$, so we can apply Theorem 5, the path coupling theorem of Dyer and Greenhill.
The yields:

$$\mathbb{E}(d(h^+, g^+)) \leq \beta \cdot d(h, g)$$

for all $(h, g) \in \Omega \times \Omega$, where (h^+, g^+) is a pair of random variables, distributed according to the application of one step of P to (h, g). This means, the application of Theorem 4 is possible. It yields

$$\tau_{M_{\mathcal{B}}}(\varepsilon) \leq \frac{\log(D \cdot \frac{1}{\varepsilon})}{1 - \beta}.$$

The last step of this proof is the translation of this bound into a bound of $\tau_M(\varepsilon)$, the mixing time of our up/down Markov chain. This is achieved with the help of Theorem 6. It needs a bound to A, which was defined as

$$A := \max_{(x,y) \in E(M)} \left\{ \frac{1}{\pi(x) \cdot M(x,y)} \cdot \sum_{(u,v) \in \Gamma(x,y)} |\gamma_{uv}| \cdot \pi(u) \cdot \widetilde{M}(u,v) \right\},$$

where $\gamma_{u,v}$ is a path of transitions of M for every transition $u \mapsto v$ of $M_{\mathcal{B}}$ and $\Gamma(x, y)$ is the set of all these paths, which use the transition $x \mapsto y$ of M.

We do not define these paths explicitly, but $\Omega_{B|l}$ is a distributive lattice of k-heights, so we can apply Lemma 2 and know, that there is a set of paths for every $u, v \in \Omega_{B|l}$ such that $|\gamma_{u,v}| \leq k \cdot |B| \leq k \cdot b$ with $b := \max\{|B| : B \in \mathcal{B}\}$.
Further, we have $M(x,y) = \frac{1}{4 \cdot |F(G)|}$ for $x \neq y$, where $M(x,y)$ is the transition probability from x to y in one step of M.
Now, for a fixed transition $x \mapsto y$ of M let f be the face which is touched. Now we have

$$\begin{aligned}
A_{xy} &:= \frac{1}{\pi(x)M(x,y)} \sum_{(u,v)\in\Gamma(x,y)} |\gamma_{u,v}| \cdot \pi(u) \cdot M_{\mathcal{B}}(u,v) \\
&= \frac{1}{M(x,y)} \sum_{(u,v)\in\Gamma(x,y)} |\gamma_{u,v}| \cdot M_{\mathcal{B}}(u,v) \\
&\leq 4 \cdot |F(G)| \sum_{(u,v)\in\Gamma(x,y)} k \cdot b \cdot M_{\mathcal{B}}(u,v) \\
&\leq 4 \cdot k \cdot b \cdot |F(G)| \sum_{\{B\in\mathcal{B}|f\in B\}} \sum_{u\in\Omega_B} \underbrace{\sum_{v\in\Omega_{B|u}} M_{\mathcal{B}}(u,v)}_{=1} \\
&= 4k \cdot b \cdot |F(G)| \sum_{B\in\mathcal{B}, f\in B} |\Omega_B| \\
&\leq 4k \cdot b \cdot |F(G)| \sum_{B\in\mathcal{B}, f\in B} (k+1)^b \\
&\leq 4k \cdot (k+1)^b \cdot b \cdot m \cdot |F(G)|,
\end{aligned}$$

which directly yields $A = \max\{A_{xy} \mid (x,y) \in E(M)\} \leq 4k \cdot (k+1)^b \cdot b \cdot m \cdot |F(G)|$. Further, we have $\frac{1}{\pi^\star} = |\Omega| \leq (k+1)^{|F(G)|}$. Now Theorem 6 yields:

$$\begin{aligned}
\tau_M(\varepsilon) &\leq \frac{4\log(\frac{1}{\varepsilon \cdot \pi_\star})}{\log(\frac{1}{2\varepsilon})} \cdot A \cdot \tau_{M_{\mathcal{B}}}(\varepsilon) \\
&\leq \frac{4\log(\frac{1}{\varepsilon} \cdot (k+1)^{|F(G)|})}{\log(\frac{1}{2\varepsilon})} \cdot (4k \cdot (k+1)^b \cdot b \cdot m \cdot |F(G)|) \cdot \frac{\log(D \cdot \frac{1}{\varepsilon})}{(1-\beta)} \\
&\leq \underbrace{\frac{16 \cdot b \cdot m \cdot k \cdot (k+1)^b}{1-\beta}}_{=c_{\mathcal{B},k}} \cdot \frac{(\log(\frac{1}{\varepsilon}) \cdot |F(G)| + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k \cdot F(G)}{\varepsilon})}{\log(\frac{1}{2\varepsilon})},
\end{aligned}$$

which is, what was claimed in the first hand. □

Let us close this section with a corollary of Theorem 2. It simplifies the condition of β, but makes it harder to find such a β:

Corollary 1
Let $\mathcal{B}$ be a set of blocks, such that every face $\delta \in F(G)$ occurs in m blocks and in $\leq \ell$ borders of blocks. Let $E := \max_{B\in\mathcal{B},\delta\in\partial B} E_{B,\delta}$ and $\beta \in \mathbb{R}$ such that

$$1 > \beta \geq 1 + \frac{1}{2|\mathcal{B}|}(\ell \cdot E - m).$$

This implies

$$\tau_M(\varepsilon) \le c_{\mathcal{B},k} \cdot \frac{(\log(\frac{1}{\varepsilon}) \cdot |F(G)| + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})},$$

where $b := \max_{B \in \mathcal{B}}\{|B|\}$ *and* $c_{\mathcal{B},k} := \frac{16 \cdot b \cdot m \cdot k \cdot (k+1)^b}{1-\beta}$, *i.e. Theorem 2 is applicable.*

Proof. We have

$$\begin{aligned}
\beta &\ge 1 + \frac{1}{2|\mathcal{B}|}(\ell \cdot E - m) \\
&\ge 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B \in \mathcal{B} | \delta \in \partial B} E \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right) \\
&\ge 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B \in \mathcal{B} | \delta \in \partial B} E_{B,\delta} \quad - \#\{B \in \mathcal{B} \mid \delta \in B\}\right)
\end{aligned}$$

So we can apply Theorem 2 and get the desired bound to the mixing time $\tau_M(\varepsilon)$ of the up/down Markov chain M. □

2.3 Toroidal triangle grid graphs and triangle grid graphs

Let us start with a simple, straight forward application of Corollary 1.
To do so, we introduce *toroidal triangle grid graphs.*
Starting with the usual, infinite triangle grid with vertices $\mathbb{Z}^2$ and edges

$$\{((x,y),(x,y+1)), ((x,y),(x+1,y)), ((x,y),(x+1,y+1)) \mid x,y \in \mathbb{Z}\},$$

we extract a finite rhombus and glue its upper border to the lower one and its left border to the right one, as shown in Figure 2.8.
This results in a non planar graph G, which covers a torus with a triangle grid. So we call G a *toroidal triangle grid graph.* Every vertex of G has degree 6, so G is regular. Further, all faces are triangles and the simple structure if G allows us to create a set of Blocks $\mathcal{B}$, which cover the faces of G nicely:

Theorem 7
Let G be a toroidal triangle grid graph of size n and let Ω be the set of its 2*-heights. The up/down Markov chain M is rapidly mixing on G, i.e.*

$$\tau_M(\varepsilon) \le 699\ 840 \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}.$$

Proof. We apply block coupling as introduced earlier. For every vertex $v \in V(G)$, we add one block B_v to the family of blocks $\mathcal{B}$. This block B_v consists of the six faces, which touch v, i.e. v is on the border of them.
This ensures, that every face f is contained in exactly $m = 3$ blocks, since every face contains three vertices v_1, v_2, v_3 such that $f \in B_{v_1}$, $f \in B_{v_2}$, and $f \in B_{v_3}$. Further, f is

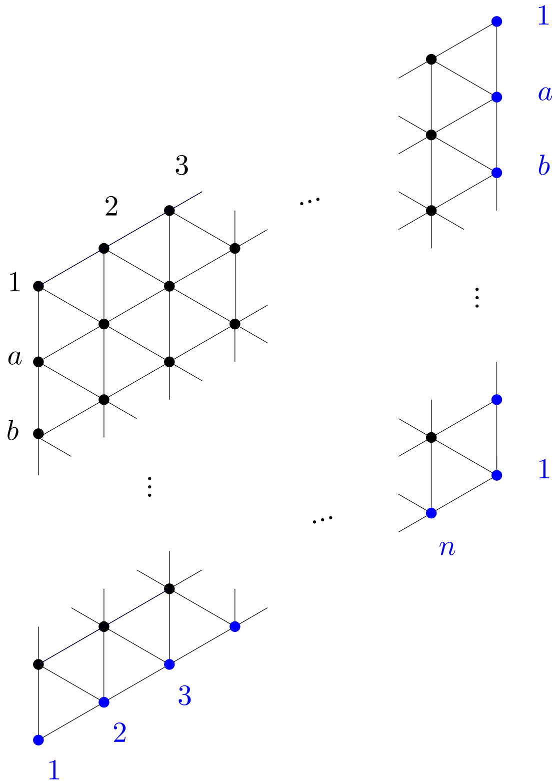

Figure 2.8: A rhombic section of the triangular grid. identifying blue and black vertices with the same name gives a toroidal triangle grid graph.

adjacent to three other faces, namely f_1, f_2, and f_3. Every block B, such that $f \in \partial B$ has to contain one of them. But each f_i shares two of its three vertices with f. So only the third vertex w_i creates a block B_{w_i}, such that $f_i \in B_{w_i}$ and $f \in \partial B_{w_i}$. This means, that for every face f there are exactly $\ell = 3$ blocks B such that f is on the border of B.

To estimate $E_{B,\delta}$, note, that due to symmetry it does neither depend on δ, nor on B. All blocks B are shifted δ the same graph. This graph has a lot of symmetries, containing rotations, which permit to rotate a fixed face δ to every of the six faces on the border of B. In total there are 729 border conditions, or elements in $\Omega_{\partial B}$. They result in 486 cover relations (u, l), or pairs of border conditions, which differ in δ.
By complete enumeration with a small c++ program, we computed

$$\mathbb{E}(\omega(u_B^+)) = \frac{1}{|\Omega_{B|u}|} \sum_{u_B^+ \in \Omega_{B|u}} \omega(u_B^+) = \frac{1}{|\Omega_{B|u}|} \sum_{u_B^+ \in \Omega_{B|u}} \sum_{f \in B} u_B^+(f)$$

for all border conditions u, which implies $\mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+))$ for every cover relation (l, u). By doing so, we learned

$$\max_{(u,l)} \left(\mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+)) \right) \approx 0.79866.$$

Double counting the pairs of faces and blocks, which contain each other leads to $6 \cdot |\mathcal{B}| = 3 \cdot |F(G)|$. So we have $E_{B,\delta} \leq 0.8$ and $|\mathcal{B}| = \frac{1}{2} \cdot |F(G)|$. Therefore:

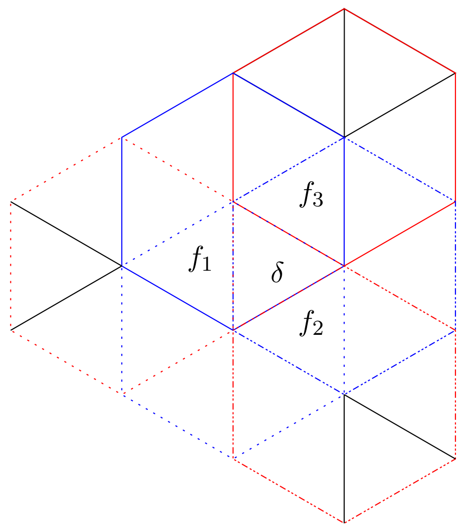

Figure 2.9: All six blocks, which contain δ (blue), or contain δ on its border (red).

$$\begin{aligned}
1 + \frac{1}{2|\mathcal{B}|}(\ell \cdot E - m) &= 1 + \frac{1}{2 \cdot \frac{1}{2} \cdot |F(G)|}(3 \cdot E - 3) \\
&\leq 1 + \frac{1}{|F(G)|}(3 \cdot 0.8 - 3) \\
&\leq 1 - \frac{6}{10 \cdot |F(G)|} \\
&=: \beta < 1
\end{aligned}$$

Now we can apply Corollary 1. Using $b = \max\{|B| : B \in \mathcal{B}\} = 6$, this leads to

$$\begin{aligned}
c_{\mathcal{B},k} &= \frac{16 \cdot b \cdot m \cdot k \cdot (k+1)^b}{1-\beta} \\
&= \frac{16 \cdot 6 \cdot 3 \cdot 2 \cdot (3)^6}{\frac{6}{10 \cdot |F(G)|}}, \\
&= 699\ 840 \cdot |F(G)|
\end{aligned}$$

and therefore implies

$$\begin{aligned}
\tau_M(\varepsilon) &\leq c_{\mathcal{B},k} \cdot \frac{(\log(\frac{1}{\varepsilon}) \cdot |F(G)| + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \\
&= 699\ 840 \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \\
&\in \mathcal{O}\left(\frac{|F(G)|^3 \cdot \log(|F(G)|)}{\log(\frac{1}{2\varepsilon})} + |F(G)|^3 + |F(G)|^2 \cdot \log(\tfrac{1}{\varepsilon}) \right),
\end{aligned}$$

which proves, what was claimed. □

Besides this theorem, we also considered other subsections of the triangle lattice. They are covered by the following theorem:

Corollary 2
Let G be a finite subset of the triangle grid graph. Let ∂G be the border of G and b a 2-height on ∂G. The up/down Markov chain M is rapidly mixing on Ω_G and $\Omega_{G|b}$, i.e.

$$\tau_M(\varepsilon) \leq 699\ 840 \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}$$

in both cases.

Proof. To prove this theorem, we use the same family of blocks $\mathcal{B}$, which we already applied in Theorem 7. So for every vertex $v \in V(G)$ we have one block B_v, which consists of all faces in G, which have v on the border. Opposing to Theorem 7, we now have different shapes of blocks. Especially, if the vertex v is on the border of G, the resulting block is new.

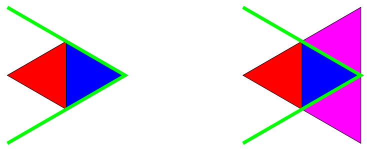

Figure 2.10: Blocks of type 1 and 1b with 1 blue face, 1 red face on the border (and 2 magenta external border constraints).

Figure 2.11 up to Figure 2.14 show these truncated sixgons. The faces in ∂B are indicated in magenta, so b might create additional border constraints for the case of $\Omega_{G|b}$, but δ can never be one of the magenta faces. Further, the edges/ border between G and ∂G is indicated in green, to clarify the situation.

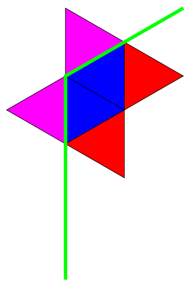

Figure 2.11: Blocks of type 2 and 2b with 2 blue faces, 2 red faces on the border (and 2 magenta external border constraints).

Note, that there are many more cases, namely for any presented case, every subset of red faces might be outside of G. This obviously does not matter for the case of $\Omega_{G|B}$, because if $E_{B,\delta}$ is maximal for a face δ, which is in fact not element of ∂B, but an external restraint (i.e. magenta instead of red), the computed 'wrong' E is bigger than the actual one, so if we can apply our machinery to the computed E, we are safe.

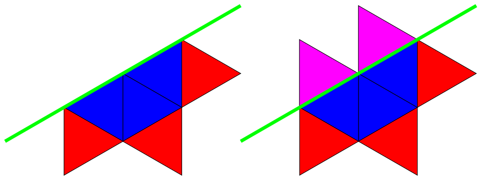

Figure 2.12: Blocks of type 3 and 3b with 3 blue faces, 3 red faces on the border (and 2 magenta external border constraints).

But it does not even matter for the case of Ω_G. First one might think, that the additional red conditions, might result in a smaller E, but a value of 1 for the faces, which are red but not present, does not influence Ω_G here, because we are considering 2-heights, where every possible value accepts the value 1 as a neighbour. So for the special case of 2-heights, which we consider, the additional red faces do not matter.
For all these type of blocks B, we estimated E_B by complete enumeration with c++ programs. The resulting values of are presented in Table 2.1.

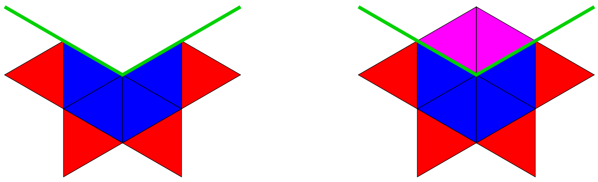

Figure 2.13: Blocks of type 4 and 4b with 4 blue faces, 4 red faces on the border (and 2 magenta external border constraints).

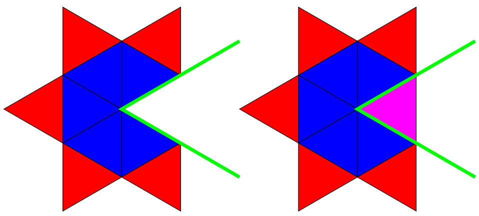

Figure 2.14: Blocks of type 5 and 5b with 5 blue faces, 5 red faces on the border (and 1 magenta external border constraint).

As one can easily check, all values of E_B are < 0.8. Further, as in the proof of Theorem 7, every face is contained in exactly three blocks and is on the border of at most three blocks. This means, we can define β as we did before as

$$\beta := 1 - \frac{6}{10 \cdot |F(G)|}$$

without external constraints

Type	E	$\lvert B\rvert/\lvert\partial B\rvert$
1	0.5	1/1
2	0.6	2/2
3	0.667	3/3
4	0.742	4/4
5	0.765	5/5

with external constraints

Type	E	$\lvert B\rvert/\lvert\partial B\rvert/$ external constraints
1b	0.5	1/1/2
2b	0.6	2/2/2
3b	0.667	3/3/2
4b	0.742	4/4/2
5b	0.765	5/5/1

Table 2.1: E_B of truncated sixgons for $k = 2$. Although values are identical here, they might in general differ.

and end up with the mixing time

$$\tau_M(\varepsilon) \le 699\ 840 \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}.$$

□

Now we know, that for the context of 2-heights on (toroidal) triangle grid graphs, everything is fine, but sadly, we are not able to present results for bigger k, because all kinds of blocks which we tried (and were able to compute) failed.

2.4 Toroidal rectangular grid graphs

This section covers the rectangular lattice. To be more specific, we consider finite, toroidal sections of the rectangular lattice. Let us start with a simple, straight forward application of Corollary 1. To do so, we introduce *toroidal rectangular grid graphs*.
As done in the previous section done with the triangle grid to create toroidal triangle grids, we start with the usual, infinite rectangular grid. This means we have vertices $\mathbb{Z}^2$ and edges

$$\{((x,y),(x,y+1)),((x,y),(x+1,y)) \mid x,y \in \mathbb{Z}\},$$

we extract a finite rhombus and glue its upper border to the lower one and its left border to the right one, as shown in Figure 2.15.

Theorem 8 *(2-heights on toroidal rectangular grid graphs mix rapidly)*
The up/down Markov chain M rapidly mixes the set of 2-heights of toroidal rectangular lattice grid graphs G, i.e.

$$\tau_M(\varepsilon) \le 512 \cdot 3^{40} \frac{|F(G)|^2 \cdot (\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{1.44 \cdot \log(\frac{1}{2\varepsilon})}.$$

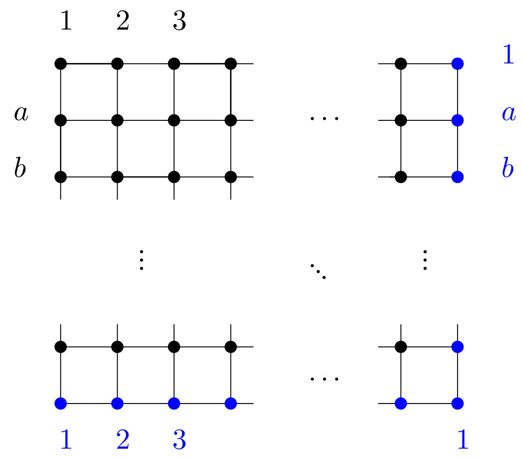

Figure 2.15: A rectangular section of the rectangle grid. identifying blue and black vertices with the same name gives a toroidal rectangular grid graph.

Proving this theorem was much harder than expected. We finally succeeded with blocks, which are squares of 6×6 faces. To compute the expected distance $E_B = \max_{\delta \in \partial B} E_{B,\delta}$ for these blocks, straightforward complete enumeration of all cases did not work, because there are too many cases, to handle without using highly optimized c++ programs and a cluster, which we did. Due to the level of optimization, the software was not able to handle any other case any more, so it was not possible to check, whether it returned the right result. Therefore, we made use of transfer matrices, to bundle a lot of cases together.

To do so, we separate ∂B into the sides S, the lower border L and upper border U (more details are given in the next definition). The goal here is, to compute for every fixed k-height on the sides $s \in \Omega_S$ two matrices $\hat{M} \in \mathbb{R}^{|\Omega_L| \times |\Omega_U|}$ and $\hat{W} \in \mathbb{R}^{|\Omega_L| \times |\Omega_U|}$, such that for an upper border-constraint $\nu \in \Omega_U$ and a lower border-constraint $\ell \in \Omega_L$ we have $\hat{M}_{\ell,\nu} = |\Omega_{B|b}|$ and $\hat{W}_{\ell,\nu} = \sum_{h \in \Omega_{B|b}} \omega(h)$, where b is the total border constraint for B, consisting of s, ν, and ℓ.
This means, instead of computing all k-heights for all border conditions, we now have to compute two matrices, for each k-height on the side, which is a subset of the border. If the matrices can be computed efficiently and are not to big, this speeds up the process a lot (for the case of 6×6 squares, the computation time changed from two weeks on a cluster to 1.5 hours on one computer, but it depends on the exact structure of the blocks to consider).
These matrices $\hat{W}$ and $\hat{M}$ are created layer by layer, using a transfer matrix method. This allows us to compute the expected height $\mathbb{E}(\omega(h))$ of the set of k–heights $\Omega_{B|b}$ much faster than the recursive method, because we do not have to create every k-height one by one. On the other hand, it needs (in some cases) a lot more memory, because all k-heights are considered at the same time, using matrix operations. The method we use here is adapted from transfer matrices as Stanley introduces them in chapter 4.7 of his book [46].
So let us explain, how to create $\hat{M}$ and $\hat{W}$:

Definition 11 (layer decomposition)
Let B be a block. A *layer decomposition* $L_0, L_1, \dots, L_m$ of B is a decomposition of the faces B, such that the *layers* L_i form a partition of B and

$$\partial\left(\bigcup_{i<j} L_i\right) \subseteq L_j,$$

where $\partial M := \{f \in F(G) \setminus M \mid \text{ there is } g \in M, \text{ sharing an edge with } f\}$ is the *border* of M. Further, the *upper border* $U \subseteq \partial B$ is the set of face of ∂B, which share an edge with a face of L_0, the *sides* $S \subseteq (\partial B \setminus U)$ are the faces of ∂B, that share an edge with a face of $L_1 \cup \ldots \cup L_{m-1}$ (and are not part of the upper border). Finally, the lower border $L \subseteq (\partial B \setminus (U \cup S))$ is the set of faces of ∂B, which are not part of the upper border or the sides. This implies, they share an edge with a face of L_m!

Next, we say a *good layer decomposition* of a block B is a layer decomposition, such that no face of the upper border U shares an edge with a face of $L_1 \cup \ldots \cup L_m$ and no face of the sides S shares an edge with a face of L_m. Further, we want B to be symmetric, such that we can assume, $\delta \in U$.

Finally, we say, that an lower border-restraint $\ell \in \Omega_L$ *fits* to the sides $s \in \Omega_S$, if there is a border-constraint $b \in \Omega_{\partial B}$, which is equal to s on the sides and equal to ℓ on the lower border and an upper border constraint ν *fits* to the sides $s \in \Omega_S$, if it suffices to corresponding condition.

Note, that every layer decomposition of G is a valid m–height and each m–height is a valid layer–decomposition, if you assign every face in layer i the value i (or vice versa).

Definition 12 (compatible k-heights)
With Ω_{L_i} being the set of all k-heights on L_i, we say $h \in \Omega_{L_i}$ *is compatible with* $g \in \Omega_{L_{i+1}}$ if and only if $h \cup g \in \Omega_{L_i \cup L_{i+1}}$, which is defined by

$$(h \cup g)(f) := \begin{cases} h(f) & \text{if } f \in L_i \\ g(x) & \text{if } f \in L_{i+1} \end{cases}.$$

This means $h \cup g$ is a valid k-height on $L_i \cup L_{i+1}$. Further, the matrix $T_i \in \{0,1\}^{\Omega_{L_{i+1}} \times \Omega L_i}$ with

$$(T_i)_{h,g} = \begin{cases} 1 & \text{if } h \in \Omega_{L_{i+1}} \text{ is compatible with } g \in \Omega_{L_i} \\ 0 & \text{else} \end{cases}$$

is the *layer transfer matrix* from layer i to layer $i+1$. Further, for every layer i we have a *layer compatibility matrix* $C_i \in \{0,1\}^{\Omega_{L_i} \times \Omega_{L_i}}$, which is a diagonal matrix with entries

$$C_i(b)_{h,h} = \begin{cases} 1 \text{ if } h \in \Omega_{L_i|b} \\ 0 \text{ else} \end{cases}.$$

So $C_i(b)$ keeps out all entries of k-heights on L_i, which are compatible to the border constraint b and maps all other entries to 0.

Note, that for a good layer decomposition, $C_0(b)$ only depends on the upper border U, $C_1(b)$ up to $C_{m-1}(b)$ only depend on the sides S and $C_m(b)$ only depends on the lower

border L, because the faces of the corresponding layers only share edges with faces of the corresponding parts of the border!
Now, the crucial observation is the following one:

Lemma 6 *(counting with transfer matrices)*
Let B be a block and $L_0, \ldots, L_m$ be a layer decomposition of B. Consider

$$v := T_{j-1} \cdot \ldots \cdot T_1 \cdot T_0 \cdot \mathbb{1}_{|\Omega_{L_0}|} \in \mathbb{N}^{|\Omega_{L_j}|},$$

where $\mathbb{1}_{|\Omega_{L_0}|} \in \mathbb{N}^{|\Omega_{L_0}|}$ is the all-one-vector. For $h \in \Omega_{L_j}$, there are v_h different heights $g \in \Omega_{L_0 \cup \ldots \cup L_j}$ with $g_{|L_j} = h$, i.e. $g(f) = h(f)$ for all $f \in \Omega_{L_j}$.

Further, let $b \in \Omega_{\partial B}$ be a border condition of B. Consider

$$w := (C_j(b) \cdot T_{j-1}) \cdot C_{j-1}(b) \cdot \ldots \cdot (C_2(b) \cdot T_1) \cdot (C_1(b) \cdot T_0) \cdot (C_0(b) \cdot \mathbb{1}_{|\Omega_{L_0}|}) \in \mathbb{N}^{|\Omega_{L_j}|}.$$

For $h \in \Omega_{L_j|b}$, there are w_h different heights $g \in \Omega_{(L_0 \cup \ldots \cup L_j)|b}$ with $g_{|L_j} = h$. In the case of $h \in \Omega_{L_j} \setminus \Omega_{L_j|b}$ we have $w_h = 0$.

Proof. We prove the first claim inductively. For $j = 1$ the formula reduces to $\mathbb{1}_{|\Omega_{L_0}|}$, which is for sure correct. Now consider $v := T_{j-1} \cdot \ldots \cdot T_1 \cdot T_0 \cdot \mathbb{1}_{|\Omega_{L_0}|}$. By induction hypothesis we know there are v_h heights on $L_0 \cup \ldots \cup L_j$ equal h if restricted to L_j. We can continue each of this heights on layer $j+1$ with the heights $\{k \in H_{j+1} \mid (T_j \cdot e_i)_k = 1\}$ by definition of T_j. Summing over all i yields:

$$\begin{aligned}
\sum_{i \in H_j} T_j \cdot e_i \cdot v_i &= \sum_{i \in H_j} T_j \cdot e_i \cdot T_{j-1} \cdot \ldots \cdot T_0 \cdot \mathbb{1}_{|\Omega_{L_0}|} \\
&= T_j \cdot \left(\sum_{i \in H_j} e_i \right) \cdot T_{j-1} \cdot \ldots \cdot T_0 \cdot \mathbb{1}_{|\Omega_{L_0}|} \\
&= T_j \cdot T_{j-1} \cdot \ldots \cdot T_0 \cdot \mathbb{1}_{|\Omega_{L_0}|}
\end{aligned}$$

To prove the second claim, use the same arguments as above. For the case $j = 0$, we have $w = C_0(b) \cdot \mathbb{1}_{|\Omega_{l_0}|}$ which is by definition the characteristic function of $\Omega_{L_0|b}$ in Ω_{L_0}. Inductively as above, $T_j \cdot (C_j(b) \cdot T_{j-1}) \cdot \ldots \cdot (C_1(b) \cdot T_0) \cdot (C_0(b) \cdot \mathbb{1}_{|\Omega_{L_0}|})$ counts all k-heights $\Omega_{L_j \cup ((L_{j-1} \cup \ldots \cup L_0)|b)}$, which are compatible with b on layer 0 up to $j-1$. This implies, $(C_j(b) \cdot T_j) \cdot (C_j(b) \cdot T_{j-1}) \cdot \ldots \cdot T_0 \cdot (C_0(b) \cdot \mathbb{1}_{|\Omega_{L_0}|})$ makes them compatible on all layers and ensures $w_h = 0$ for all $h \in \Omega_{L_j} \setminus \Omega_{L_j|b}$. □

Corollary 3
Let B be a block with layer decomposition $L_0, \ldots, L_m$ and $b \in \Omega_{\partial B}$ a border constraint. Now with

$$M := T_{m-1} \cdot (C_{m-1}(b) \cdot T_{m-2}) \cdot \ldots \cdot (C_2(b) \cdot T_1) \cdot (C_1(b) \cdot T_0)$$

and

$$W := \sum_{j=0}^{m-1} T_{m-1} \cdot (C_{m-1}(b) \cdot T_{m-2}) \cdot \ldots \cdot (T_j \cdot C_j(b)) \cdot \omega_j \cdot (T_{j-1} \cdot C_{j-1}(b)) \cdot \ldots \cdot (C_1(b) \cdot T_0)$$

we have

$$|\Omega_{B|b}| = \mathbb{1}^T_{|\Omega_{L_m}|} \cdot C_m(b) \cdot M \cdot C_0(b) \cdot \mathbb{1}_{|\Omega_{L_0}|}$$

and

$$\mathbb{E}(\omega(h)) = \frac{1}{|\Omega_{B|b}|}\mathbb{1}^T_{|\Omega_{L_m}|}\cdot C_m(b)\cdot W\cdot C_0(b)\cdot \mathbb{1}_{|\Omega_{L_0}|},$$

for a random variable h, living on $\Omega_{B|b}$ uniformly distributed. Here $\omega_j \in \mathbb{N}^{|\Omega_{L_j}|\times|\Omega_{L_j}|}$ is the diagonal matrix, with entries $(\omega_j)_{h,h} = \omega(h) = \sum_{f\in L_j} h(f)$, i.e. it weights a vector of k-heights on L_j according to their weight.

Proof. To prove the first claim, apply Lemma 6 and sum up over all entries of v.

To argue for the second claim, consider the contribution c_j of layer L_j to $\mathbb{E}(\omega(h))\cdot|\Omega_{B|b}|$. This means

$$c_j = \sum_{h\in\Omega_{B|b}}\sum_{f\in L_j} h(f)$$

and therefore

$$\mathbb{E}(\omega(h))\cdot|\Omega_{B|b}| = \sum_{h\in\Omega_{B|b}}\omega(h) = \sum_{h\in\Omega_{B|b}}\sum_{f\in B} h(f) = \sum_{j=0}^{m} c_j$$

Now, according to Lemma 6 by applying it twice, we have

$$\left(\mathbb{1}^T_{|\Omega_{L_m}|b|}\cdot(C_m(b)\cdot T_{m-1})\cdot\ldots\cdot C_j(b)\cdot e_h\right)\cdot\left(e_h^T\cdot C_j(b)\cdot\ldots\cdot T_0\cdot C_0(b)\cdot\mathbb{1}_{|\Omega_{L_0}|b|}\right)$$

heights $g\in\Omega_{B|b}$,which are equal to $h\in\Omega_{L_j|b}$ on all faces of layer j.
Summing over all $h\in\Omega_{L_j|b}$ yields

$$\begin{aligned}
c_j &= \sum_{h\in\Omega_{L_j}}\mathbb{1}^T_{|\Omega_{L_m}|b|}\cdot(C_m(b)\cdot T_{m-1})\cdot\ldots\cdot C_j(b)\cdot e_h\cdot \underbrace{\omega(h)}_{=(\omega_j)_{h,h}}\cdot e_h^T\cdot C_j(b)\cdot\ldots\cdot T_0\cdot C_0(b)\cdot\mathbb{1}_{|\Omega_{L_0}|b|}\\
&= \mathbb{1}^T_{|\Omega_{L_m}|b|}\cdot(C_m(b)\cdot T_{m-1})\cdot\ldots\cdot C_j(b)\cdot\omega_j\cdot C_j(b)\cdot T_{j-1}\cdot\ldots\cdot T_0\cdot C_0(b)\cdot\mathbb{1}_{|\Omega_{L_0}|b|}\\
&= \mathbb{1}^T_{|\Omega_{L_m}|b|}\cdot(C_m(b)\cdot T_{m-1})\cdot\ldots\cdot C_j(b)\cdot\omega_j\cdot T_{j-1}\cdot\ldots\cdot T_0\cdot C_0(b)\cdot\mathbb{1}_{|\Omega_{L_0}|b|},
\end{aligned}$$

because ω_j and $C_j(b)$ commute (they are both diagonal matrices) and $C_j(b)^2 = C_j(b)$ ($C_j(b)$ is a $\{0,1\}$- diagonal matrix). Now, summing up over all columns j proves to the claim. ☐

To add even some more notation, let $T_U\in\mathbb{R}^{|\Omega_{L_0}|\times|\Omega_U|}$ and $T_L\in\mathbb{R}^{|\Omega_L|\times|\Omega_{L_m}|}$ be matrices with entries

$$\begin{aligned}
(T_U)_{h,\nu} &:= \begin{cases}1 & \text{if } h\in\Omega_{L_0|\nu}\\ 0 & \text{else}\end{cases} \quad\text{and}\\
(T_L)_{\ell,h} &:= \begin{cases}1 & \text{if } h\in\Omega_{L_m|\ell}\\ 0 & \text{else}\end{cases}
\end{aligned}$$

This means $T_U\cdot e_b = C_0(b)\cdot\mathbb{1}_{|\Omega_{L_0}}$ and $e_b^T\cdot T_L = \mathbb{1}_{|\Omega_{L_m}}\cdot C_m(b)$, where e_b is the b-th canonical unit vector, implicitly restricting b to the upper (or lower) border. So we have

$$|\Omega_{B|b}| = \mathbb{1}^T_{|\Omega_{L_m|b}|}\cdot M\cdot\mathbb{1}_{|\Omega_{L_0|b}|} = (T_L\cdot M\cdot T_U)_{b_{|L},b_{|U}}$$

and

$$\mathbb{E}(\omega(g)) = \frac{1}{|\Omega_{B|b}|} \mathbb{1}^T_{|\Omega_{L_m|b}|} \cdot W \cdot \mathbb{1}_{|\Omega_{L_0|b}|} = \frac{(T_L \cdot W \cdot T_U)_{b_{|L}, b_{|U}}}{(T_L \cdot M \cdot T_U)_{b_{|L}, b_{|U}}},$$

for a random variable g, uniformly distributed on $\Omega_{B|b}$.

To conclude, with $\hat{M} = T_L \cdot M \cdot T_U$ and $\hat{W} = T_L \cdot W \cdot T_U$ we found he two promised matrices. Next, Algorithm 4 tells us, how to estimate $E_{B,\delta}$, using these two matrices.

Input: block B with good layer decomposition $L_1, \dots, L_m$
Output: $(E_{B,\delta} \mid \delta \in U)$
$E \longleftarrow (0, 0, \dots, 0) \in \mathbb{R}^{|U|}$
$M \longleftarrow T_U$
$W \longleftarrow \omega_1 \cdot T_U$
foreach $s \in \Omega_S$ **do**
    for $j = 2$ *to* m **do**
        $H \longleftarrow C_j(s) \cdot T_{j-1}$
        $M \longleftarrow H \cdot M$
        $W \longleftarrow \omega_j \cdot M + H \cdot W$
    $\hat{M} \longleftarrow T_L \cdot M$
    $\hat{W} \longleftarrow T_L \cdot W$
    foreach (l, u) *cover relation of* Ω_U*, which fits to the sides* s **do**
        $\delta \longleftarrow$ the face $\in U$, u and l differ in
        $h \longleftarrow \max\left\{\left|\frac{\hat{W}_{u,\ell}}{\hat{M}_{\ell,u}} - \frac{\hat{W}_{\ell,l}}{\hat{M}_{\ell,l}}\right| : \ell \in \Omega_L \text{ fitting to } s\right\}$
        if $h > E_\delta$ **then**
            $E_\delta \longleftarrow h$
return E

Algorithm 4: The transfer matrix method to estimate $E_{B,\delta}$ for all $\delta \in U$.

It is fairly easy to see, that this algorithm indeed returns the correct result, because by construction we have

$$h = \left|\frac{\hat{W}_{\ell,u}}{\hat{M}_{\ell,u}} - \frac{\hat{W}_{\ell,l}}{\hat{M}_{\ell,l}}\right| = |\mathbb{E}(\omega(u_B^+)) - \mathbb{E}(\omega(l_B^+))|,$$

where u_B^+ is a random variable, uniformly distributed on $\Omega_{B|u}$ and l_B^+ is a random variable, uniformly distributed on $\Omega_{B|l}$. So the maximum of this term over all cover relations $(l, u) \in \Omega_{\partial B}$, which differ in δ is $E_{B,\delta}$. Further we have $\{(s, \nu, \ell) \mid s \in \Omega_S, \nu \in \Omega_U, \ell \in \Omega_L \text{ and } \nu, \ell \text{ fit to } s\}$ is isomorphic to $\Omega_{\partial B}$. The algorithm computes the maximum of h on this set, so it estimates $E_{B,\delta}$. Let us consider this method on a small example:

Example 3
Consider the block B, which is displayed in Figure 2.16 and assume $k = 2$. The block B consists of four blue faces and has six faces on its border. The numbers in the blue faces indicate the layer, each face is assigned to, so L_0 consists of one face, L_2 of two, and L_3 consists again of one face. Therefore, the upper border consists of all faces, which are

adjacent to the blue face, labeled with 0. So $U = (u_0, u_1)$. In the same manner we get sides $S = (s_0, s_1)$ and the lower border $L = (l_0, l_1)$.

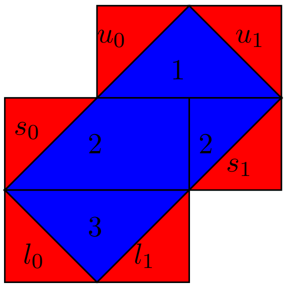

Figure 2.16: The block B of Example 3.

This means, we have

$$\begin{aligned}\Omega_{L_1} &= ((0),(1),(2)),\\ \Omega_{L_2} &= ((0,0),(0,1),(1,0),(1,1),(1,2),(2,1),(2,2)),\\ \Omega_{L_3} &= ((0),(1),(2)), \text{ and}\\ \Omega_U = \Omega_L &= ((0,0),(0,1),(0,2),(1,0),(1,1),(1,2),(2,0),(2,1),(2,2)).\end{aligned}$$

Next,we can establish all transfer matrices, which are introduced in Definition 12. We have

$$T_1 = \begin{pmatrix} 1 & 1 & 0\\ 1 & 1 & 0\\ 1 & 1 & 0\\ 1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 1 & 1\\ 0 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{|\Omega_{L_2}|\times|\Omega_{L_1}|} = \mathbb{R}^{7\times 3},$$

because $(T_1)_{h,g} = 1$ if $h \in \Omega_{L_2}$ is compatible with $g \in \Omega_{L_1}$ and 0 otherwise by definition. Further, we get

$$T_2 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0\\ 1 & 1 & 1 & 1 & 1 & 1 & 1\\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{|\Omega_{L_3}|\times|\Omega_{L_2}|} = \mathbb{R}^{3\times 7}.$$

Now, assume, we want to compute the expected height of B for the border b, defined by $u_0 = 0$, $u_1 = 2$, $s_0 = 1$, $s_1 = 2$, $l_0 = 1$, and $l_1 = 0$. We can easily check by hand, that only eight such 2-heights exists. They are $(1,0,1,0)$, $(1,0,1,1)$, $(1,1,1,0)$, $(1,1,1,1)$, $(1,1,2,0)$, $(1,1,2,1)$,$(1,2,1,1)$ and $(1,2,2,1)$. So we know

$$\mathbb{E}(\omega(h)) = \frac{1}{|\Omega_{B|b}|} \sum_{h\in\Omega_{B|b}} \sum_{f\in B} h(f) = \frac{32}{8} = 4.$$

Applying Algorithm 4, which should return the same result, we next compute the matrices M, W so we need T_U, $C_2(b)$, and T_L, which were defined in Corollary 3.

We have

$$T_U = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{|\Omega_{L_1}|\times|\Omega_U|} = \mathbb{R}^{3\times 9},$$

and again, due to symmetry $T_L = T_U^T \in \mathbb{R}^{9\times 3}$. Remember, the entries of T_U were defined in the same manner as the entries of the transfer matrices, so an entry $(T_U)_{h,u} = 1$ if $h \in \Omega_{L_1}$ is compatible to $u \in \Omega_U$ and 0 otherwise. Next, note that $C_2(b)$ only depends on the entries of the sides S, which are adjacent to to L_2. This means, we have

$$C_2((1,2)) = \operatorname{diag}(0,1,0,1,1,1,1) \in \mathbb{R}^{|\Omega_{L_2}|\times|\Omega_{L_2}|} = \mathbb{R}^{7\times 7},$$

so the diagonal entries of all 2-heights of L_2, which are compatible with the subset $s_0 = 1$ and $s_2 = 2$ of the sides S are 1 and all other entries are 0 (in this example are the complete sides, but usually, it is even only a subset of S). With weight matrices $\omega_1 = \operatorname{diag}(0,1,2)$, $\omega_2 = \operatorname{diag}(0,1,1,2,3,3,4)$, and $\omega_3 = \operatorname{diag}(0,1,2)$, where the entry $(\omega_i)_{h,h} = \sum_{f\in\Omega_{L_i}} h(f)$ is the weight of the 2-height $h \in \Omega_{L_i}$. Now we can perform the needed matrix multiplications and get

$$\begin{aligned} \hat{M} &= T_L \cdot M \cdot T_U \\ &= T_L \cdot T_2 \cdot C_2((1,2)) \cdot T_1 \cdot T_U \\ &= \begin{pmatrix} 12 & 12 & 8 & 12 & 18 & 14 & 8 & 14 & 14 \\ 12 & 12 & 8 & 12 & 18 & 14 & 8 & 14 & 14 \\ 7 & 7 & 5 & 7 & 11 & 9 & 5 & 9 & 9 \\ 12 & 12 & 8 & 12 & 18 & 14 & 8 & 14 & 14 \\ 17 & 17 & 12 & 17 & 27 & 22 & 12 & 22 & 22 \\ 12 & 12 & 9 & 12 & 20 & 17 & 9 & 17 & 17 \\ 7 & 7 & 5 & 7 & 11 & 9 & 5 & 9 & 9 \\ 12 & 12 & 9 & 12 & 20 & 17 & 9 & 17 & 17 \\ 12 & 12 & 9 & 12 & 20 & 17 & 9 & 17 & 17 \end{pmatrix} \text{ and} \\ \hat{W} &= T_L \cdot W \cdot T_U \\ &= T_L \cdot \big(\omega_3 \cdot T_2 \cdot C_2((1,2)) \cdot T_1 + T_2 \cdot C_2((1,2)) \cdot \omega_2 \cdot T_1 + T_2 \cdot C_2((1,2)) \cdot T_1 \cdot \omega_1\big) \cdot T_U \\ &= \begin{pmatrix} 40 & 40 & 32 & 40 & 73 & 65 & 32 & 65 & 65 \\ 40 & 40 & 32 & 40 & 73 & 65 & 32 & 65 & 65 \\ 28 & 28 & 23 & 28 & 52 & 47 & 23 & 47 & 47 \\ 40 & 40 & 32 & 40 & 73 & 65 & 32 & 65 & 65 \\ 68 & 68 & 56 & 68 & 129 & 117 & 56 & 117 & 117 \\ 56 & 56 & 47 & 56 & 108 & 99 & 47 & 99 & 99 \\ 28 & 28 & 23 & 28 & 52 & 47 & 23 & 47 & 47 \\ 56 & 56 & 47 & 56 & 108 & 99 & 47 & 99 & 99 \\ 56 & 56 & 47 & 56 & 108 & 99 & 47 & 99 & 99 \end{pmatrix} \end{aligned}$$

and see that indeed $\frac{\hat{W}_{(1,0),(0,2)}}{\hat{M}_{(1,0),(0,2)}} = \frac{32}{8} = 4$, as we already knew. Note, that with this matrix multiplications we did not only compute the expected height for this one border conditions, but for all border conditions $\Omega_{\partial B|S}$, which have sides $s_0 = 1$ and $s_1 = 2$. So by symmetry of B we can assume $\delta \in L$. This means, we only have to iterate over all possible sides, computing the corresponding matrices $\hat{M}$ and $\hat{W}$ and are able to compute all values of $E_{B,\delta}$, without listing all 2-heights (or k-heights in a more general setting) of Ω_B explicitly. This method enabled us, to speed up the computation of $E_{B,\delta}$, such that we were able to prove Theorem 8:

Proof. (of Theorem 8)
The set of blocks, which we use to prove this theorem, consists of shifts of squares, each consisting of 6×6 faces. We have to compute $E_B = \max_{\delta \in \partial B} E_{B,\delta}$ for these blocks, but they all have the same structure. This means E_B does not depend on the shift, so we have to compute only one value.
We use the transfer matrix approach, so there is the need of a layer decomposition of this block. Here, each layer is one row of the square, starting with the first row as layer L_0, as indicated in Figure 2.17. This means, we have six layers, each consisting of six faces. Further the upper and lower border each consists of six faces and the sides contain twelve faces. Due to the symmetries of the square, we can assume $\delta \in \{u_0, u_1, u_2\}$, so is obvious, that this is indeed a good layer decomposition.

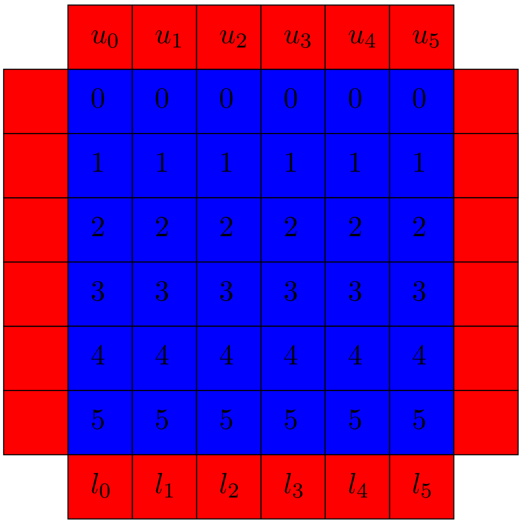

Figure 2.17: the square of 6×6 faces with red border. Numbers indicate, in which layer the face is. $\{u_0, \ldots, u_5\}$ is the upper border and $\{l_0, \ldots, l_5\}$ the lower one.

Applying a Matlab implementation of Algorithm 4 to this good layer decomposition, we get a vector $(E_{B,u_0}, E_{B,u_1}, E_{B,u_2}) = (1.2366, 1.4035, 1.4312)$, so $E = \max_{B \in \mathcal{B}} \max_{\delta \in \partial B} E_{B,\delta} = 1.4312$.
Now, every face $f \in F(G)$ occurs in $m = 36$ blocks B and is on the border of $\ell = 24$ blocks, which yields

$$\begin{aligned} 1 + \frac{1}{2|\mathcal{B}|}(\ell \cdot E - m) &\leq 1 + \frac{1}{2|\mathcal{B}|}(24 \cdot 1.44 - 36) \\ 1 - \frac{1.44}{2|\mathcal{B}|} &=: \beta < 1, \end{aligned}$$

so we can apply Corollary 1. With $|\mathcal{B}| = |F(G)|$ and $b = \max_{B \in \mathcal{B}} |B| = 36$ we get

$$\begin{aligned} c_{\mathcal{B},k} &= \frac{16 \cdot b \cdot m \cdot k \cdot (k+1)^b}{1-\beta} \\ &= \frac{16 \cdot 36 \cdot 36 \cdot 2 \cdot (2+1)^{36}}{\frac{1.44}{2|\mathcal{B}|}} \\ &= \frac{512 \cdot 3^{40} |F(G)|}{1.44} \end{aligned}$$

and

$$\begin{aligned} \tau_M(\varepsilon) &\leq c_{\mathcal{B},k} \cdot \frac{(\log(\frac{1}{\varepsilon}) \cdot |F(G)| + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \\ &= 512 \cdot 3^{40} \cdot \frac{|F(G)|^2 \cdot (\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{1.44 \cdot \log(\frac{1}{2\varepsilon})} \\ &\approx 4.32 \cdot 10^{21} \cdot \frac{|F(G)|^2 \cdot (\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \\ &\in \mathcal{O}\left(\frac{|F(G)|^3 \log(\frac{2 \cdot |F(G)|}{\varepsilon}) + |F(G)|^2 \cdot \log(\frac{1}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \right) \end{aligned}$$

□

Note that there is no simple analogue to Corollary 2, which showed that the up/down Markov chain M rapidly mixes set of 2-heights of arbitrary sections of the triangle grid. In that Corollary we considered 'truncated sixgons'. Here we would have to consider all truncated 6×6 squares, which are a lot cases to consider, especially, since there are a lot of faces, which are covered by different sets of truncated squares, which all would have to be considered separately.
In the same way as in the previous section, we obviously tried to lift this result from $k = 2$ to $k = 3$. But again, we failed in the same way due to lack of computational power. Therefore, it is time to close this section and consider different families of graphs.

2.5 Plane Triangulations

This section covers one result for outerplane triangulations. These are embedded planar graphs, such that every face, including the outer one is a triangle. So from now on, assume that G is a connected outer plane triangulation with at least 5 vertices and a minimal degree ≥ 2. For $k = 2$ we are now able to present a positive result:

Theorem 9 *(2-heights on plane outer triangulations mix rapid)*
For $G \in \mathcal{F}_\Delta$, the up/down Markov chain rapidly mixes the set of 2-heights Ω of G, i.e.

$$\tau_M(\varepsilon) \leq 43\ 082\ 150\ 400 \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}$$

Proof. To prove this theorem, we apply Theorem 2, so we need a set of blocks $\mathcal{B}$. In fact, we use two kind of blocks. For every vertex $v \in V(G)$, which has a degree of 10 or smaller, we have 8 identical blocks, which consists of all faces $f \in F(G)$, which are touching the vertex v. These are the blocks of *first type*. For every vertex $v \in V(G)$, which has a degree $\deg(v) > 10$, we have $\deg(v)$ different blocks of *second type*, which each consist of eight consecutive faces, $f_1, \ldots, f_8$, which touch v, so every face next to v has the role of f_1 in one of these blocks. One of these blocks is shown in Figure 2.18.

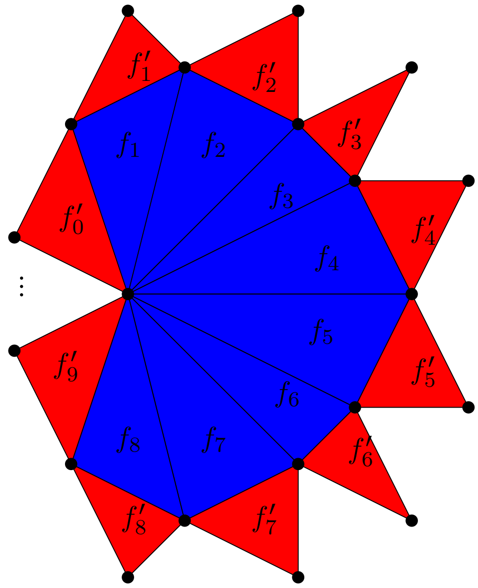

Figure 2.18: A block of second type, consisting of eight blue faces with ten red faces on the border.

First, we assume, that no face $\delta \in F(G)$ appears twice on the border of a block B. This means, the face δ shares never an edge with two distinct faces $f_1, f_2 \in B$. With this assumption, we have only nine different types of blocks left. So we can check each of them. A computer program did this by performing a complete enumeration. Its results are listed in Table 2.2. Note, that in any case we have $E_{B_v,\delta} \leq 0.8$.

$\deg(v)$	B	$\lvert\partial B\rvert$	$E_{B_v,\delta}$
2	2	2	0.6
3	3	3	0.72
4	4	4	0.77
5	5	5	0.79
6	6	6	0.80
7	7	7	0.80
8	8	8	0.80
9	9	9	0.80
≥ 10	8	10	0.79

Table 2.2: E_{B_v} of the blocks, induced by a vertex v

Further, every face $f \in F(G)$ is contained in exactly $m = 3 \cdot 8$ blocks (eight for each of the

three vertices, which it touches) and it is in $\ell \leq 30$ blocks an element of the border. At most 24 of these blocks result from three third vertices v_1, v_2, v_3 of the faces f_1, f_2, and f_3, which share an edge (and two vertices) with δ. Each of this blocks has $E_{B,\delta} \leq 0.8$, according to Table 2.2. Additionally, there might be six more blocks, which contain δ on the border, because each of the three vertices of δ can have a degree ≥ 10, which would mean, that for every such vertex two additional blocks were added to $\mathcal{B}$, which have δ on its border. Each of these blocks has $E_{B,\delta} \leq 0.79$.
This implies

$$\sum_{B \in \mathcal{B} | \delta \in \partial B} E_{B,\delta} \leq 24 \cdot 0.80 + 6 \cdot 0.79 = 23.94$$

which leads to

$$\begin{aligned} & 1 + \frac{1}{2|\mathcal{B}|} \left(\sum_{B \in \mathcal{B} | \delta \in \partial B} - \#\{B \in \mathcal{B} \mid \delta \in B\} \right) \\ \leq \quad & 1 + \frac{1}{2|\mathcal{B}|} (23.94 - 24) \\ = \quad & 1 - \frac{0.06}{2 \cdot 8 \cdot |F(G)|} \\ = \quad & 1 - \frac{3}{800 \cdot |F(G)|} \\ =: \quad & \beta < 1. \end{aligned}$$

Further, we have $b = \max_{B \in \mathcal{B}} \{|B|\} \leq 9$ and $|\mathcal{B}| = 8 \cdot |V(G)|$ and we can apply Theorem 2. Doing so, we get

$$\begin{aligned} c_{\mathcal{B},k} &= \frac{16 \cdot b \cdot m \cdot k \cdot (k+1)^b}{1 - \beta} \\ &= \frac{16 \cdot 9 \cdot 24 \cdot 2 \cdot (3)^9}{\frac{3}{800 \cdot |F(G)|}} \\ &= 16 \cdot 9 \cdot 24 \cdot 2 \cdot 3^8 \cdot 800 \cdot |F(G)| \\ &= 43\ 082\ 150\ 400 \cdot |F(G)|. \end{aligned}$$

This finally leads to

$$\begin{aligned} \tau_M(\varepsilon) \quad &\leq \quad c_{\mathcal{B},k} \cdot \frac{(\log(\frac{1}{\varepsilon}) \cdot |F(G)| + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \\ &= \quad 43\ 082\ 150\ 400 \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot |F(G)|}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}, \end{aligned}$$

which is the claim of this theorem.

Before closing this proof, we still have to prove the assumption. Remember, we assumed, that it does not matter if a face δ shares an edge with two faces f_1, f_2 of one block B. To formalize this, let δ be such a face, as shown in Figure 2.19.

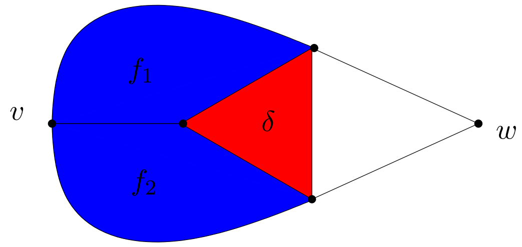

Figure 2.19: A face δ, such that two neighbouring faces f_1 and f_2 share their third vertex v.

We will show

$$\sum_{B \in \mathcal{B} | \delta \in \partial B} E_{B,\delta} \leq 23.94,$$

which is, what we have for all faces, which do not appear twice in the border of a block. We think, there is an elegant argument to show this, based on the structure of the coupling on paths, as described in the second theorem of Dyer and Greenhill (Theorem 5) and the idea, to split δ into two faces. This would lead to a pair of 2-heights, which differ in two faces, i.e. have a distance of 2. Still, the technical details of this proof seem to be tricky. We simply compute $E_{B,\delta}$ of the corresponding blocks B with computer programs. In total, we have to check twelve cases. Eight of them result in the merging of two faces on the border of a block, induced by a vertex of degree < 10 - one case, for every degree. One case suffices, because these blocks are symmetric with respect to rotations, so it does not matter, which of the faces are merged. For vertices of degree ≥ 10, we have 4 cases, because –although the block structure does not depend on $\deg(v)$– we do not have a rotation symmetry here. We only can mirror these blocks, so it suffices, to merge the faces f'_k and f'_{k+1} (according to the notation of Figure 2.18) for $k = 1, 2, 3, 4$. Note, that f'_0 and f'_1 cannot be merged, because this would imply a double edge between v and the second vertex, which f_1 and f_2 share. The resulting values of E_{B_v} are printed in Table 2.3. Note, that all of them are ≤ 1.60.

$\deg(v)$	E_{B_v}
2	0.5
3	1
4	1.22
5	1.33
6	1.38
7	1.4
8	1.41
9	1.41

$\deg(v)$	k	E_{B_v}
≥ 10	0	1.41
≥ 10	1	1.41
≥ 10	2	1.41
≥ 10	3	1.41

Table 2.3: E_{B_v} of the blocks, induced by a vertex v, where a face δ appears twice on the border

To finally prove the assumption, we have at most eight blocks B_v, which have δ twice on its border, which leads to $E_{B_v,\delta} \leq 1.6$. Further, there are at most eight blocks, which are induced by the vertex w and have δ on its border and $E_{B_w,\delta} \leq 0.8$. Finally, there are at most six blocks B, which which contain δ on the border, because each of the three vertices of δ has a degree ≥ 10. They have $E_{B,\delta} \leq 0.79$. Together, this yields

$$\sum_{B\in\mathcal{B}|\delta\in\partial B} E_{B,\delta} \leq 8 \cdot 1.41 + 8 \cdot 0.8 + 6 \cdot 0.79 \leq 23.94,$$

which is, what we wanted to show. So the assumption is true, which finishes this proof. □

Remark 2
One might ask, what happens of δ shares all three edges with faces f_1, f_2, and f_3, which share their third vertex. If that is the case, as depicted in Figure 2.20, the graph G is either not connected or contains only 4 vertices, which we both excluded in the beginning.

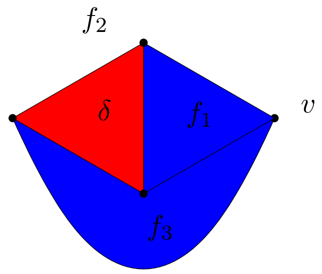

Figure 2.20: A face δ, such that all three neighbouring faces share their third vertex v. We omitted to paint the outer face f_2 in blue, although it is in the block B_v.

Corollary 4 *(2-heights on plane outer triangulations of degree ≤ 10 mix rapidly)*
Let G be a plane outer triangulation with maximum degree $d < 10$. The mixing time $\tau_M(\varepsilon)$ of the up/down Markov chain M mixing the set of 2-heights of G is bounded by:

$$\tau_M(\varepsilon) \leq 2\ 880 \cdot 3^d \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot F(G)}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}$$

Proof. Here, no vertex v has a degree of 11 or bigger, so we can completely omit the blocks of second type. This also means, that we only have to insert one block of the first type for every vertex $v \in V(G)$! Thereby, we know that every face is contained in three blocks and is in the border of tree blocks. Plugging these values, combined with the values of $E_B \leq 0.8$ of Table 2.2 into Corollary 1 yields:

$$\sum_{B\in\mathcal{B}|\delta\in\partial B} E_{B,\delta} \leq 3 \cdot 0.80 = 2.40$$

This implies

$$\begin{aligned}
& 1 + \frac{1}{2|\mathcal{B}|}\left(\sum_{B\in\mathcal{B}|\delta\in\partial B} -\#\{B \in \mathcal{B} \mid \delta \in B\}\right) \\
\leq\ & 1 + \frac{1}{2|\mathcal{B}|}(2.4 - 3) \\
=\ & 1 - \frac{0.6}{2 \cdot |F(G)|} \\
=\ & 1 - \frac{3}{10 \cdot |F(G)|} \\
=:\ & \beta < 1.
\end{aligned}$$

Further, we have $b = \max_{B\in\mathcal{B}}\{|B|\} \leq d$ and $|\mathcal{B}| = |V(G)|$ and we can apply Theorem 2. Doing so, we get

$$\begin{aligned}
c_{\mathcal{B},k} &= \frac{16 \cdot b \cdot m \cdot k \cdot (k+1)^d}{1-\beta} \\
&= \frac{16 \cdot 9 \cdot 3 \cdot 2 \cdot (3)^d}{\frac{3}{10 \cdot |F(G)|}} \\
&= 16 \cdot 9 \cdot 2 \cdot 3^d \cdot 10 \cdot |F(G)| \\
&= 2\ 880 \cdot 3^d \cdot |F(G)|
\end{aligned}$$

which finally leads to

$$\begin{aligned}
\tau_M(\varepsilon) &\leq c_{\mathcal{B},k} \cdot \frac{(\log(\frac{1}{\varepsilon}) \cdot |F(G)| + |F(G)|^2 \cdot \log(k+1)) \cdot \log(\frac{k \cdot F(G)}{\varepsilon})}{\log(\frac{1}{2\varepsilon})} \\
&= 2\ 880 \cdot 3^d \cdot |F(G)|^2 \cdot \frac{(\log(\frac{1}{\varepsilon}) + |F(G)| \cdot \log(3)) \cdot \log(\frac{2 \cdot F(G)}{\varepsilon})}{\log(\frac{1}{2\varepsilon})},
\end{aligned}$$

□

2.6 Conclusion

Before closing this chapter, we would like to pose some open questions, which are more or less obvious. First, the constants in the bounds to the mixing times are really big. Is it possible to improve them significantly? The bound for the parameter A of the comparison theorem of Randal and Tetali 6 might be improved for general graphs, or even for the special cases, considered here. Especially the term $(k+1)^b$, where $b = \max\{|B| : B \in \mathcal{B}\}$ appears to be to generous.

The next question covers bigger k. For which families of graphs is the up/down Markov chain M rapidly mixing for $k > 2$? And are there blocks, which suffice to show this? In general what are 'good' blocks, i.e. blocks which measure up to our conditions? On the one hand, a block B with a small border and a lot of faces on the inside seems to be good, because a bigger $E_{B,\delta}$ is possible for the block. On the other hand, a lot of faces on the border mean a lot of influence from the border onto the inner faces, so the influence of the one face δ is / might be smaller, which would minimize $E_{B,\delta}$. Are there other families of graphs, to which we can apply this machinery? For example does it work an plane quadrangulations? And finally, to which other settings is the block coupling technique applicable?

3 Lattice Path Enumeration

As we have seen in the previous chapter, k-heights seem to be rather reluctant against block coupling, at least for bigger k. To go further then $k = 2$, we tried to understand the structure of k-heights in another context and started to consider the one-dimensional case. These one-dimensional k-heights can be understood as lattice paths and lattice paths are, what these chapter covers. Sadly, we did not make it back from lattice paths to k-heights, but we assume, this excursion is interesting on its own.

3.1 Preliminaries

Let $P_n(\alpha)$ be the path of $n+1$ vertices with a loop of weight α at each vertex and path-edges of weight 1, see Figure 3.1. We label the vertices from one end to the other, starting with vertex 0 and ending with vertex n. A walk in a graph is a list $x_0, e_1, x_1, \ldots, e_m, x_m$ of vertices x_i and edges e_j, such that edge e_i has endpoints x_{i-1} and x_i. A walk is oriented, it starts in x_0 and ends in x_m, we say that the walk is from x_0 to x_m. The length of a of a path is the number of edges traversed, it is m in the example. With a walk in an edge-weighted graph we define its weight as the product of the weights of its edges. For example: $1 \to 2 \to 3 \to 3 \to 2$ is a walk from 1 to 2 in $P_5(\alpha)$. Its length is 4 and its weight is $1 \cdot 1 \cdot \alpha \cdot 1 = \alpha$.

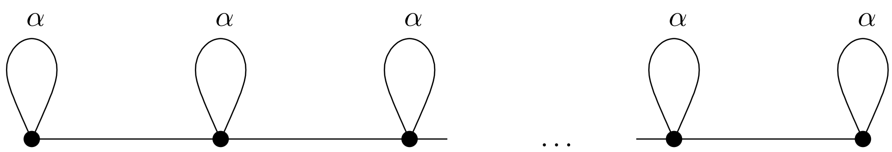

Figure 3.1: The path $P_n(\alpha)$ with $n+1$ vertices and loops of weight α.

Counting walks on $P_n(\alpha)$ enables us to count Motzkin paths in a strip: Motzkin paths of length n are lattice paths with steps of three types $(x,y) \to (x+1, y+1)$, $(x,y) \to (x+1, y)$, and $(x,y) \to (x+1, y-1)$ that start in $(0,0)$ end in $(n,0)$ and stay above the x-axis, i.e, $y \geq 0$. A Motzkin path is in the strip $[0,k]$ iff its y-coordinate never exceeds k.

strip, i.e., Motzkin paths which do not exceed a given level, see Figure 3.2 (Motzkin paths are lattice paths with steps $i \to i+1$, $i \to i$, and $i+1 \to i$ and $i \geq 0$).

Other types of lattice walks in a strip (for example Dyck paths or Schröder paths) can also be interpreted as walks on paths, see Section 3.3. We use the weighted adjacency matrix of $P_n(\alpha)$ as a tool to enumerate classes of such lattice paths. In principle, this is nothing but a variant of the traditional transfer-matrix method. Results similar to ours but tailored towards an audience of physicists and with a focus on random walks have been obtained by by Cicuta et al. [11].

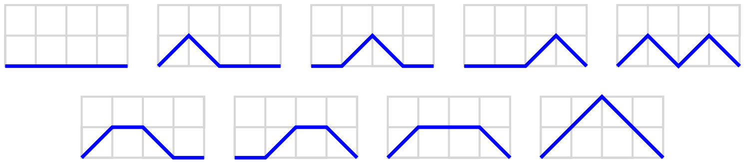

Figure 3.2: The 9 Motzkin paths of length 4; only the last one exceeds the first level.

The enumeration of classes of lattice paths is a classical topic originating from Bertrand's ballot problem. They are investigated in probability theory in the context of gamblers ruin problem (see [19], chapter XIV) but also in their own right. The monograph by Mohanty [40] still gives a valuable overview. A recent survey on lattice path enumeration is due to Humphreys [32].

3.2 The counting approach

Let us start with two simple facts about the enumeration of walks in graphs.

Fact 1
Let G be a graph with weights on the edges and let A be its weighted adjacency matrix. The sum of the weights of walks of length m from vertex i to vertex j is

$$e_i^T \cdot A^m \cdot e_j.$$

From now on we focus on the graph $P_n = P_n(1)$ and its weighted version $P_n(\alpha)$ respectively. With $A_{(n,\alpha)}$ we denote the weighted adjacency matrix of $P_n(\alpha)$; clearly $A_{(n,\alpha)} \in \mathbb{R}^{(n+1)\times(n+1)}$.

Definition 13
$Z^n_{i,j}(m,\ell) := \#(\text{walks from } i \text{ to } j \text{ of length } m, \text{ using } \ell \text{ loops in } P_n).$

The following fact is obtained from Fact 1 and simple combinatorial reasoning.

Fact 2

$$Z^n_{i,j}(m,\ell) = \binom{m}{\ell} \cdot Z^n_{i,j}(m-\ell,0) \tag{3.1}$$

$$\sum_{\ell=0}^{m} Z^n_{i,j}(m,\ell)\ \alpha^\ell = e_i^T \cdot A^m_{(n,\alpha)} \cdot e_j \tag{3.2}$$

These facts can be deduced by simple combinatoric means, respectively rely on the meaning of weighted powers of adjacency matrices, as stated in Fact 1. Extending these facts, the following theorem states the key identities for the numbers $Z^n_{i,j}(m,\ell)$. In particular this leads to explicit expressions for these numbers in terms of trigonometric sums.

Theorem 10

$$Z^n_{i,j}(m-\ell,0) = \frac{2}{n+2}\sum_{k=1}^{n+1}\left(2\cdot\cos\left(\frac{k\cdot\pi}{n+2}\right)\right)^{m-\ell}\sin\left(i\cdot\frac{k\cdot\pi}{n+2}\right)\sin\left(j\cdot\frac{k\cdot\pi}{n+2}\right) \tag{3.3}$$

$$e_i^T\cdot A^m_{(n,\alpha)}\cdot e_j = \frac{2}{n+2}\sum_{k=1}^{n+1}\left(\alpha+2\cos\left(\frac{k\cdot\pi}{n+2}\right)\right)^{m}\sin\left(i\cdot\frac{k\cdot\pi}{n+2}\right)\sin\left(j\cdot\frac{k\cdot\pi}{n+2}\right) \tag{3.4}$$

To prove the two formulas given in the theorem we begin with simple linear algebra. The first proposition is well known:

Proposition 2 *Let $A \in \mathbb{R}^{(n+1)\times(n+1)}$ be a matrix and assume that A admits an orthonormal basis $(v_1,\dots,v_{n+1})$ of eigenvectors. Also let $\lambda_i \in \mathbb{C}$ be the eigenvalue corresponding to $v_i = (v_{1,i},\dots,v_{n+1,i})$. Then*

$$e_i^T \cdot A^m \cdot e_j = \sum_{k=1}^{n+1} \lambda_k^m v_{i,k} v_{j,k}.$$

Proof. Consider the matrix $S = [v_1,\dots,v_{n+1}]$, i.e., $S\cdot e_i = v_i$ for all i. From orthogonality we obtain $S^T = S^{-1}$. This implies $e_i = S^T \cdot v_i$ and $S^T \cdot A \cdot S = \mathrm{diag}(\lambda_1,\dots,\lambda_{n+1}) =: D$. Now $A^m = (S\cdot D\cdot S^{-1})^m = S\cdot D^m \cdot S^{-1} = S\cdot D^m\cdot S^T$ and hence, $e_i^T\cdot A^m\cdot e_j = e_i^T S\cdot D^m\cdot S^T e_j = (S^T e_i)^T\cdot D^m\cdot(S^T e_j) = (v_{i,1},\dots,v_{i,n+1})^T\cdot D^m\cdot(v_{j,1},\dots,v_{j,n+1}) = \sum_{k=1}^{n+1}\lambda_k^m v_{i,k}v_{j,k}$. □

$A_{(n,\alpha)}$ is a real symmetric matrix. So it has a real orthonormal basis $(v_1,\dots,v_{n+1}) \in \mathbb{R}^{n+1}$ of eigenvectors. The next lemma relates eigenvalues and eigenvectors of $A_{(n,\alpha)}$ and of $A_{(n,0)}$.

Lemma 7 *Let $\lambda_1,\dots,\lambda_{n+1}$ be the eigenvalues of $A_{(n,0)}$ and $(v_1\dots,v_{n+1})$ a corresponding orthonormal basis of eigenvectors.Therefore all eigenvalues of $A_{(n,\alpha)}$ are of the form $\lambda_i+\alpha$ and $(v_1,\dots,v_{n+1})$ is an orthonormal basis of eigenvectors of $A_{(n,\alpha)}$.*

Proof. Let v be an eigenvector of $A_{(n,0)}$ with respect to λ, which implies $A_{(n,\alpha)}\cdot v = (A_{(n,0)}+\alpha\cdot I_n)\cdot v = A_{(n,0)}\cdot v+\alpha\cdot v = (\lambda+\alpha)\cdot v$. Hence, v is an eigenvector of $A_{(n,\alpha)}$ w.r.t the eigenvalue $\lambda+\alpha$. □

Our next aim is to determine the orthonormal basis of eigenvectors of $A_{(n,0)}$. The structure of this matrix is captured by the next definition.

Definition 14 (Toeplitz matrix)
A matrix $A = (a_{ij}) \in \mathbb{R}^{(n+1)\times(n+1)}$ is called a *Toeplitz matrix* if there exist $c_{-n}, c_{-n+1},\dots, c_0,\dots c_n \in \mathbb{R}$ such that

$$a_{ij} = c_{i-j} \text{ for all } i,j = 1,2,\dots,n+1$$

i.e. A is constant on all diagonals. If $c_{-k} = c_k$ the matrix A is a *symmetric Toeplitz matrix.* We will denote A as $\mathtt{toep}(c_{-n},\dots,c_n)$ to have a short notation for A.

Now, the adjacency matrix of $P_n(\alpha)$ is $A_{(n,\alpha)} = \mathtt{toep}(0,\dots,0,1,\alpha,1,0,\dots,0)$.

Proposition 3 *The eigenvalues λ_k and (right) eigenvectors v_k of $A_{(n,0)}$ are*

$$\lambda_k := 2\cdot\cos\left(\frac{k\cdot\pi}{n+2}\right)$$

and

$$v_k := \left(\sin\left(1\cdot\frac{k\cdot\pi}{n+2}\right), \sin\left(2\cdot\frac{k\cdot\pi}{n+2}\right),\dots,\sin\left((n+1)\cdot\frac{k\cdot\pi}{n+2}\right)\right)$$

for $k = 1,\dots,n+1$. Moreover, $\|v_k\|_2^2 = \frac{n+2}{2}$ for all k.

Proof. Since $A_{(n,0)}$ is a tridiagonal Toeplitz matrix with diagonals $1, 0$, and 1, the eigenvectors and eigenvalues are known, see e.g. [9]. Replacing (a, b, c) in the given formula on page 35 of that book by $(1, 0, 1)$ yields the result. □

Now it is easy to complete the proof of formulas (3) and (4):

Proof. (Theorem 10) For (3) first replace $Z^n_{i,j}(0, m-\ell)$ by $e_i^T \cdot A^{m-\ell}_{n,0} \cdot e_j$, a consequence of (2). The proof of the formulas now follows from a combination of Proposition 2, Lemma 7 and Proposition 3. □

Remark 3
This occurrence of Chebyshev polynomials of the second kind and their roots, which are closely related to the eigenvalues above, is well known. Flajolet counts Motzkin paths with continued fractions. Abandoning the continued fractions leads to the bounded case described here. Also the work [10] of Chow and West contains relations between the Chebyshev polynomials (and their roots) and 123-avoiding permutations, which are closely linked to Motzkin paths.

Besides the exact enumeration in terms of trigonometric sums, the numbers $Z^n_{i,j}(m, \ell)$ can also be related to linear recurrences and generating functions. An approach to obtain the generating function can be found in [46], Chapter 4, Theorem 4.2.7. A linear recurrence can be derived from the characteristic polynomial of the weighted adjacency matrix or any other polynomial, which has the weighted adjacency matrix as a root:

Fact 3
Let A be a matrix and $\chi_A(x)$ be the characteristic polynomial of A. From Cayley-Hamilton we know $\chi_{A_{(n,\alpha)}}(A_{(n,\alpha)}) = 0$. If $\chi_{A_{(n,\alpha)}}(x) = x^{n+1} + \sum_{k=0}^{n} a_k x^k$ this leads to the linear recurrence:

$$(A_{(n,\alpha)})^{n+1} = -\sum_{k=0}^{n} a_k (A_{(n,\alpha)})^k.$$

With Fact 2, (3.1) we obtain

$$(A_{(n,0)})^{n+1} = -\sum_{k=0}^{n} a_k (A_{(n,0)})^k \quad \Longrightarrow \quad Z^n_{i,j}(m, \ell) = -\sum_{k=0}^{n} a_k Z^n_{i,j}(m-n+j, \ell).$$

Besides, the characteristic polynomial of $A_{(n,\alpha)}$ is known to be determined by the $n+1$-th Chebyshev polynomial of second kind $U_{n+1}(x)$ as

$$\chi_{A_{(n,\alpha)}}(x) = U_{n+1}\left(\tfrac{x-\alpha}{2}\right).$$

From this the explicit expressions for the coefficients a_k can be obtained. This then leads to an explicit linear recurrence for $Z^n_{i,j}(m, \ell)$.

3.3 Interesting special cases

In this section we show that the numbers $Z^n_{i,j}(m, \ell)$ are quite universal. Special combinations of their parameters lead to a multitude of well known and not so well known sequences. In many cases we refer to a sequence by using its number from the *The On-Line Encyclopedia of Integer Sequences*, see [45].

Example 4 (Binomial Coefficients)

$$Z^n_{k,n-k}(n,0) = \binom{n}{k}$$

Proof. The start and end for the walks on P_n are such that any sequence of k steps down and $n-k$ steps up stays on P_n. □

The choice of the parameters is not unique. Indeed for all $k' \geq k$ and $n' \geq k' + n - k$ we have

$$Z^{n'}_{k',k'+n-2k}(n,0) = Z^{n-k}_{k,n-k}(n,0) = \binom{n}{k}.$$

Further, *bounded* or *restricted* binomial coefficients can be counted this way, where our choice of balls is restricted in such a way, that at each step of choosing our subset, the difference of the number chosen and not chosen balls is always at most k.

Example 5 (Bounded Binomial Coefficients)
Let $C(n,k,b)$ be the number of $\{0,1\}$-strings of length n with k times 1 such that for each initial segment of the string the number of 0's and the number of 1's differ at most by b. It is easy to verify that

$$C(n,k,b) = Z^{2b}_{b,2k+b-n}(n,0)$$

For $b = 1$, we have $Z^2_{1,1}(2n,0) = 2^n = Z^2_{1,2}(2n+1,0)$ and for $b = 2$ we have $Z^4_{2,2}(2n,0) = 2 \cdot 3^n$. Since the sequence $a_n := Z^6_{3,3}(2n,0)$ fulfills the recursion $a_n = 4a_{n-1} - 2a_{n-2}$ and has the same initial values it is A006012. The sequence $b_n = Z^8_{4,4}(2n,0)$ seems to be to A147748.

Example 6 (Catalan numbers)

$$Z^n_{0,0}(2n,0) = \frac{1}{n+1}\binom{2n}{n} = C_n$$

Proof. Catalan numbers count Dyck paths from $(0,0)$ to $(0,2n)$. Dyck paths correspond to walks on P_n without loops, starting at vertex 1 and returning to 1 after $2n$ steps. Traditionally, Dyck paths are described as lattice paths with steps $i \to i+1$ and $i \to i-1$ from $(0,0)$ to $(0,2n)$, which do not fall below 0. Catalan numbers count Dyck paths from $(0,0)$ to $(0,2n)$, i.e., lattice paths with steps $i \to i+1$ and $i \to i-1$ from $(0,0)$ to $(0,2n)$, which stay above the x-axis. Dyck paths correspond to walks on P_n without loops, starting at vertex 0 and returning to 0 after $2n$ steps. □

Again, this can be generalized to all $n' \geq n$ since $Z^{n'}_{0,0}(2n,0) = Z^n_{0,0}(2n,0)$. Also Dyck paths can be restricted in height.

Example 7 (Bounded Catalan numbers)
Bounded Catalan numbers in our sense count the number of Dyck paths, which do not exceed a given level k. Restricting $k = 2$, yields $Z^2_{0,0}(2m,0) = Z^2_{1,1}(2(m-1),0) = 2^{m-1}$ and for bigger k some of the sequences are well known, as the corresponding entries from the OEIS show:

$Z^3_{0,0}(2m,0)$	A001519	$Z^5_{0,0}(2m,0)$	A080937	$Z^{m-2}_{0,0}(2m,0)$	not listed
$Z^4_{0,0}(2m,0)$	A124302	$Z^6_{0,0}(2m,0)$	A024175	$Z^{m-1}_{0,0}(2m,0)$	not listed

Bounded Dyck paths have also been counted by Owczarek and Prellberg [41]. They derive a closed formula for a q-analog, where q is taking the area under the path into account.

Example 8 (Fibonacci Numbers)

$$Z^2_{0,\gcd(m,2)}(m,0) = F_m$$

Proof. F_n = number of Catalan paths from $(0,0)$ to $(n, \gcd(n,2))$ that stay between the lines $y = 0$ and $y = 3$. In the On-Line Encyclopaedia [45] this is attributed to Kimberling. □

Example 9 (Motzkin numbers)

$$\sum_{\ell=1}^{n} Z^{\lfloor n/2 \rfloor}_{0,0}(n,\ell) = M_n$$

Proof. Motzkin numbers count Motzkin paths, i.e. lattice paths from $(0,0)$ to $(0,n)$, using only steps up-left, horizontal or down-left, see Figure 3.2. These paths can be counted as walks with loops on $P_n(1)$. □

The refined counting for Motzkin numbers allows to control the maximum height of the corresponding path as well as the number of horizontal steps allowed (or at least used).

Example 10 (Bounded Motzkin numbers)
Bounded Motzkin numbers are defined similarly to bounded Catalan numbers but on the basis of Motzkin paths instead of Dyck paths.

$\sum_\ell Z^1_{0,0}(m,\ell)$	2^{m-1}	$\sum_\ell Z^5_{0,0}(m,\ell)$	A094287	$\vdots$	$\vdots$
$\sum_\ell Z^2_{0,0}(m,\ell)$	A024537	$\sum_\ell Z^6_{0,0}(m,\ell)$	A094288	$\sum_\ell Z^{\lfloor \frac{m}{2} \rfloor -1}_{0,0}(m,\ell)$	not listed
$\sum_\ell Z^3_{0,0}(m,\ell)$	A005207	$\sum_\ell Z^7_{0,0}(m,\ell)$	not listed	$\sum_\ell Z^{\lfloor \frac{m}{2} \rfloor -2}_{0,0}(m,\ell)$	not listed
$\sum_\ell Z^4_{0,0}(m,\ell)$	A094286	$\vdots$	$\vdots$		

Example 11 (Delannoy numbers)

$$\sum_{\ell=0}^{a} Z^{b+a}_{a,b-a}(a+b-\ell,\ell) = \sum_{\ell=0}^{a} \binom{a}{\ell}\binom{b+\ell}{b} = D(a,b).$$

Proof. Delannoy numbers describe the number of lattice paths from $(0,0)$ to (a,b) which use steps up, right and diagonal (i.e. steps with the effect of up+right). We do a refined counting (by counting the paths with a fixed number of diagonal steps and therefore a fixed number of total steps). The Delannoy paths from $(0,0)$ to (a,b) with $\ell \le a$ diagonal steps have length $a+b-\ell$, they use $b-\ell$ up-steps and $a-\ell$ right-steps. Modeling right-steps as decreasing-steps in the path, diagonals as loops and up-steps as increasing-steps, we see that the number of these paths is $Z^{b+a}_{a,b-a}(a+b-\ell,\ell)$. □

Example 12 (Schröder numbers)

$$\sum_{\ell=0}^{n} Z^{n}_{0,0}(2n-2\ell,\ell) = S(n).$$

Proof. Schröder numbers count lattice paths from $(0,0)$ to (n,n) with steps up, right and diagonal, which do not exceed the diagonal. Again, a counting, refined by fixing the number of diagonal steps (to fix the number of steps in the path) yields the result. All we have to change is the start and endpoint of our walks to ensure, that the path does not exceed the diagonal. □

Example 13 (Pell numbers)
Another example are Pell numbers (A000129). They are $b_n = \sum_{\ell=0}^{n} Z_{0,2}^{2}(n,\ell)$. The b_n fulfill the recursion $b_n = 2b_{n-1} + b_{n-2}$.

3.4 Refined Counting

In this section we derive linear recursions for a refined version of the ballot problem. A *turn* is a change of direction from up to down or vice versa (the up step and the down step need not be consecutive, they may enclose some loops, respectively horizontal steps). Figure 3.3 shows an example.

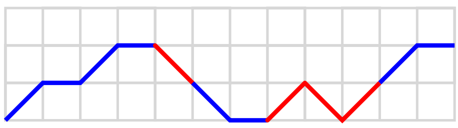

Figure 3.3: A lattice walk of length 12 with 4 red turns and 2 straight steps in $\vec{P}_3(\alpha)$ from u_0 to u_2.

Our refined counting is going to keep track of the number of *turns* of the paths. As in Section 3.2 we identify lattice walks with walks on some graph. The relevant graph $\vec{P}_n^+(\alpha)$ is obtained from the graph $\vec{P}_n(\alpha)$ shown in Figure 3.4 by adding the vertices u_0 and d_n with their loops and the edges $u_0 \to u_1$ and $d_n \to d_{n-1}$.
With a *turn* a change of direction from up to down (or vice versa) is meant. Straight steps / loops do not change the direction, so the first step down (or up) after an up-step (down-step), followed by an arbitrary number of straight ones is a turn. Figure 3.3 shows one such walk, where all turns are marked red. As it was done in Section 3.2, these lattice walks are described as walks on a graph. This time it is the graph $\vec{P}_n(\alpha)$, illustrated in Figure 3.4. .
The graph consists of two directed paths of length $n-1$ with loops which are connected by bidirectional edges in such a way, that the i-th vertex of the first path is linked to the $(n-i)$-th vertex of the second one. Weight the edges such that the loops have weight α, the edges on the first path have weight 1, the second path's edges 1 and the edges linking both paths have weight β. Note that the vertices u_0 and d_n are omitted, since they have no incoming edges and therefore do not play a role in counting the lattice walks.
Now, let

$$\begin{aligned}\vec{Z}_{v,w}^{n}(m,\ell,t) &:= \#(\text{ walks on } \vec{P}_n(\alpha) \text{ from } v \text{ to } w \text{ of length } m \text{ with } \ell \text{ loops and } t \text{ turns}),\\ \vec{Z}_{v,w}^{n}(m) &:= \sum_{\ell,t} \vec{Z}_{v,w}^{n}(m,\ell,t)\alpha^{\ell}\beta^{t}.\end{aligned}$$

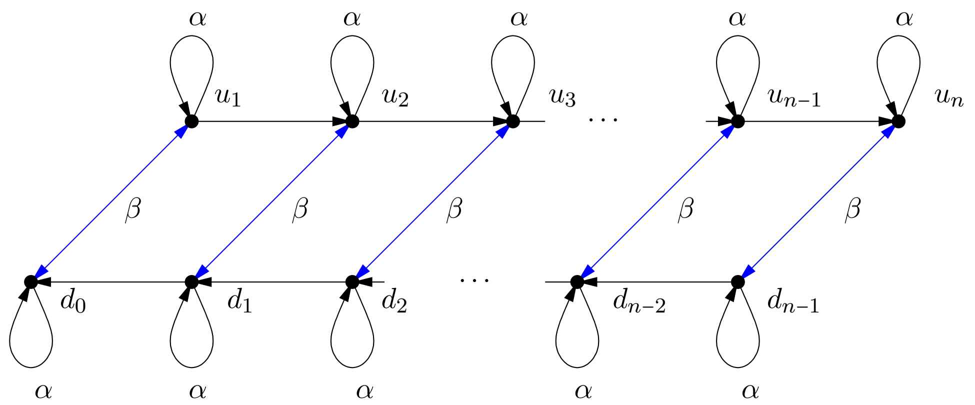

Figure 3.4: The directed path $\vec{P}_n(\alpha)$ with $2n$ vertices and loops of weight α and blue turning-edges of weight β

Krattenthaler already counted some instances of these lattice paths successfully. See [36], especially Theorem 3.4.4 for details. Still, his work does not cover straight steps/ loops, while our results do so:

Theorem 11 *(Characteristic polynomial of $\vec{P}_n(\alpha)$)*
Let $\chi_{\vec{P}_n(\alpha)}(x)$ be the characteristic polynomial of the adjacency matrix of the graph $\vec{P}_n(\alpha)$. Then:

$$\sum_{n=0}^{\infty} \chi_{\vec{P}_n(\alpha)}(x)\lambda^n = \frac{(\alpha-x)^2(\lambda-1)-\beta^2}{1+(\beta^2-1+(\alpha-x)^2(\lambda-1))\lambda}$$

and

$$\vec{Z}^n_{v,w}(m) = \frac{-1}{c^n_{2n}} \cdot \sum_{k=1}^{2n} c^n_{2n-k} \cdot \vec{Z}^n_{v,w}(m-k),$$

where

$$c^n_k = (-1)^k \sum_{j=0}^{n} \sum_{l=0}^{n-j} \sum_{i=0}^{j} \binom{n-l-j-1}{j-1}\binom{l+j}{l}\binom{j}{i}\binom{2i}{k}(-\beta^2)^{l+j-i}\alpha^{2i-k}$$

are the coefficients of the characteristic polynomials $\chi_{\vec{P}_n(\alpha)}(x)$ of the graphs $\vec{P}_n(\alpha)$.

To prove this theorem, we use a well known connection between cycle covers and determinants of the adjacency matrix of the covered (di-) graphs.

Definition 15 (Cycle cover)
Let $\vec{G}$ be a directed graph with weighted edges. A *cycle cover* C of $\vec{G}$ consists of $|V(\vec{G})|$ edges, such that every vertex has one incoming and one outgoing edge. In other words it is a set of simply directed cycles which contains every vertex exactly once.
Further, the weight $\omega(C)$ of the cycle cover C is the product of all edge-weights of C multiplied with $(-1)^{\#\text{ even cycles}}$. Finally $\mathcal{C}(\vec{G})$ is the set of all cycle covers of $\vec{G}$.

Lemma 8 *(Cycle covers and determinants)*

Let $\vec{G}$ be a digraph with edge weights and adjacency matrix $A = (a_{ij})$. We have

$$\det(A) = \sum_{\sigma \in S_n} sign(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = \sum_{C \in \mathcal{C}(\vec{G})} \omega(C) = \omega(\mathcal{C}(\vec{G}))$$

Proof. A permutation $\sigma \in S_n$ can be seen as a subset of n directed edges $C = \{(i,\sigma(i)) \mid i = 1\ldots,n\}$, which form a cycle cover. This cycle cover has weight $\mathsf{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma(i)}$. If one of the edges is not present in $\vec{G}$, the weight is 0, since the corresponding entry of A is 0. The sign of σ is the product of the signs of individual cycles. Even cycles contribute a factor of -1, odd cycles a factor $+1$. □

Next, we make use of an application of the geometric sum to matrices and generating functions:

Theorem 12 *(Theorem 4.7.2 of [46] or Proposition V.9 of [26])*
Let A be the adjacency matrix of a graph G. Further, let $N_k(i,j)$ be the number of walks from i to j of length k and $w_{ij} = \sum_{k=0}^{\infty} N_k(i,j)t^k$. This yields $W = (w_{ij}) = (\mathcal{I}_{|V(G)|} - t \cdot A)^{-1}$, where $\mathcal{I}$ is the identity matrix of the right size.

Finally, we are ready to prove the previously stated theorem:

Proof. (of Theorem 11) To find the generating function for the characteristic polynomials $\chi_{\vec{P}_n(\alpha)}$ of the adjacency matrix of the graph $\vec{P}_n(\alpha)$, we count weighted cycle covers of the graph $\vec{P}_n(\alpha - x)$. According to Lemma 8 the sum of their weights is the polynomial we are looking for. Now, we partition the set of all cycle covers, according to the edges they use to cover the vertices d_{n-1} and u_n. Let A_n be the set of all cycle covers of $\vec{P}_n(\alpha - x)$, such that the vertices u_n and d_{n-1} are both covered by loops and let B_n be the set of all remaining cycle covers. Cycle covers in B_n contain the edge (u_n, d_{n-1}), since apart from the loop this is the only outgoing edge of u_n. Further, let $a_n := \omega(A_n) = \sum_{C \in A_n} \omega(C)$ and $b_n := \omega(B_n)$. This implies $\omega(\mathcal{C}(\vec{P}_n(\alpha - x))) = \sum_{C \in \mathcal{C}(\vec{P}_n(\alpha - x))} \omega(C) = a_n + b_n$. Now, every cycle cover of $\vec{P}_{n-1}(\alpha)$ can be extended with two loops of weight $(\alpha - x)^2$ or a two-cycle of weight $-\beta^2$. Further for every cycle cover in B_{n-1}, we can replace the edge u_{n-1}, d_{n-2} by the path $u_{n-1}, u_n, d_{n-1}, d_{n-2}$. This leads to a cycle cover in B_n with the same weight. Conversely, from a cycle cover in A_n we can delete the two loops at u_n and d_{n-1} to obtain a cover of $\vec{P}_{n-1}$. In a cycle cover in B_n the two vertices u_n and d_{n-1} are either covered by a two cycle which can be deleted or they belong to a longer cycle from which they can be removed. This implies the linear recursions $a_{n+1} = (\alpha - x)^2(a_n + b_n)$ and $b_{n+1} = -\beta^2 a_n + (-\beta^2 + 1)b_n$. The initial conditions are $a_0 = (\alpha - x)^2$ and $b_0 = -\beta^2$. With $A = \begin{pmatrix} (\alpha - x)^2 & (\alpha - x)^2 \\ -\beta^2 & -\beta^2 + 1 \end{pmatrix}$, we obtain:

$$\chi_{\vec{P}_n(\alpha)}(x) = \omega(\mathcal{C}(\vec{P}_n(\alpha - x))) = a_n + b_n = (1,1) \cdot \begin{pmatrix} a_n \\ b_n \end{pmatrix} = (1,1) \cdot A \cdot \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix} = (1,1) \cdot A^n \cdot \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$$

Hence,

$$\sum_{n=0}^{\infty} \chi_{\vec{P}_n(\alpha)}(x)\lambda^n = \sum_{n=0}^{\infty} \left((1,1) \cdot A^n \cdot \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} \right) \cdot \lambda^n$$

and Theorem 12 can be applied. To do so, we need the inverse of $(\mathcal{I}_2 - \lambda A)$:

$$(\mathcal{I}_2 - \lambda \cdot A)^{-1} = \frac{1}{1 + (\beta^2 - 1)\lambda + (\alpha - x)^2(\lambda^2 - \lambda)} \begin{pmatrix} 1 - (-\beta^2 + 1)\lambda & (\alpha - x)^2\lambda \\ -\beta^2\lambda & 1 - (\alpha - x)^2\lambda \end{pmatrix}$$

Therefore, the generating function for the characteristic polynomials of $\vec{P}_n(\alpha)$ is

$$\sum_{n=0}^{\infty} \chi_{\vec{P}_n(\alpha)}(x) \cdot \lambda^n = (1,1) \cdot (\mathcal{I}_2 - \lambda A)^{-1} \cdot \begin{pmatrix} (\alpha - x)^2 \\ -\beta^2 \end{pmatrix} = \frac{(\alpha - x)^2(\lambda - 1) - \beta^2}{1 + (\beta^2 - 1 + (\alpha - x)^2(\lambda - 1))\lambda},$$

which proves the first part of the theorem. To deduce the theorem's second claim, one could establish an explicit representation of $\chi_{\vec{P}_n(\alpha)}(x)$ via a partial fraction decomposition of the generating function above, but there is a more direct and elegant approach: Consider again Lemma 8. A cycle cover of $\vec{P}_n(\alpha - x)$ decomposes the upper path of $\vec{P}_n(\alpha - x)$ into segments of consecutive vertices, each segment belonging to one cycle. Also every such decomposition implies some cycle covers, since any segment has to be closed to one cycle. For sections, which have length ≥ 1, the cycle is fixed, while for solely vertices, they can either be covered with loops or with a two-cycle.
Now, there are $\binom{n-1}{t-1}$ such decompositions into t non-empty parts (each chosen ball marks the end of an part; the last parts ends at the end, so the last vertex is chosen ahead) , but they have different weights. So they imply different numbers of cycle weights. Let us count these decompositions with respect to l, the number of sections of length 1 and j, the number of sections of length ≥ 2. This means we have to decompose $n - l$ vertices into sections j of length ≥ 2 and afterwards insert the sections of length 1 in between two sections of length ≥ 2. So there are $\binom{n-l-j-1}{j-1} \cdot \binom{j+l}{l}$ such decompositions. For each of them, the sum of the weights of the corresponding cycle covers is $(-\beta^2)^l((\alpha - x)^2 - \beta^2)^j$, as stated previously. This leads to

$$\begin{aligned}
\chi_{\vec{P}_n(\alpha)}(x) &= \sum_{j=0}^{n}\sum_{l=0}^{n-j} \binom{n-l-j-1}{j-1}\binom{l+j}{l}(-\beta^2)^l((\alpha - x)^2 - \beta^2)^j \\
&= \sum_{j=0}^{n}\sum_{l=0}^{n-j} \binom{n-l-j-1}{j-1}\binom{l+j}{l}(-\beta^2)^l \left(\sum_{i=0}^{j}\binom{j}{i}(\alpha - x)^{2i}(-\beta)^{2(j-i)}\right) \\
&= \sum_{i=0}^{n}\left(\sum_{j=i}^{n}\sum_{l=0}^{n-j} \binom{n-l-j-1}{j-1}\binom{l+j}{l}\binom{j}{i}(-\beta)^{2(j-i+l)}\right)(\alpha - x)^{2i} \\
&= \sum_{i=0}^{n}\left(\sum_{j=i}^{n}\sum_{l=0}^{n-j} \binom{n-l-j-1}{j-1}\binom{l+j}{l}\binom{j}{i}(-\beta)^{2(j-i+l)}\right)\left(\sum_{k=0}^{2i}\binom{2i}{k}\alpha^{2i-k}(-x)^k\right) \\
&= \sum_{k=0}^{2n}\underbrace{\left(\sum_{i=\lceil k/2\rceil}^{n}\sum_{j=i}^{n}\sum_{l=0}^{n-j} \binom{n-l-j-1}{j-1}\binom{l+j}{l}\binom{j}{i}\binom{2i}{k}(-\beta)^{2(j-i+l)}\alpha^{2i-k}\right)}_{=c_{k,n}}(-x)^k
\end{aligned}$$

Now, Fact 3 yields the second claim of our theorem since $\vec{Z}^n_{v,w}(m)$ is the entry of the m-th power of the adjacency matrix of $\vec{P}_n(\alpha)$, associated to v and w.

□

3.5 Further Results

3.5.1 Tri-diagonal Toeplitz matrices

In the literature you commonly find variants of the formulas in Theorem 10 with an additional parameter. This parameter represents different weights for the step $i \to i+1$ and the step $i \to i-1$. These different weights still yield a tridiagonal Toeplitz matrix. The spectrum and an orthonormal basis of eigenvectors are known for tridiagonal Toeplitz matrices, see [9]. Therefore, our theorem can easily be adapted to cover the more general case.

3.5.2 The symmetric case

For the case of symmetric lattice walks, i.e. walks on graphs, having an adjacency matrix, which is a symmetric Toeplitz matrix, we apply theorem 10 to enumerate these walks, due to the following fact:
Let $A \in \mathbb{R}^{(n+1)\times(n+1)}$ be a symmetric Toeplitz matrix and let $c_0, c_1, \dots c_n \in \mathbb{R}$ be its first row. Now A is completely determined by these values. The powers $A^k_{(n,0)}$, $k = 0, \dots, n$ of the adjacency matrix of $P_n(0)$ form a basis for symmetric Toeplitz matrices, in fact if $a^k_0, a^k_1, \dots a^k_n$ is the first row of $A^k_{(n,0)}$, implying $a^k_k = 1$ and $A^k_j = 0$ for all $j > k$.
Hence we can write

$$A = \sum_{k=0}^{n} \tilde{c_k} \cdot A^k_{(n,0)}$$

and the coefficients $\tilde{c_k}$ can be recursively computed as $\tilde{c_n} = c_n$ and $\tilde{c_k} = c_k - \sum_{j=k+1}^{n} \tilde{c}_j a^j_k$. Therefore, the orthonormal basis of eigenvectors of $A_{(n,0)}$ is an orthonormal basis of eigenvectors of A. Furthermore the eigenvalues of A are

$$\hat{\lambda_i} := \sum_{k=0}^{n-1} \tilde{c_k} \lambda_i^k,$$

where λ_i are the eigenvalues of $A_{(n,0)}$. Since we have an orthonormal basis of eigenvectors and the corresponding eigenvalues, we can write down formulas for A that correspond to the formulas (3) and (4) for $A_{(n,0)}$.

3.5.3 Lattice walks with arbitrary positive steps

For some cases case of asymmetric lattice walks, we find linear recursions. If we only allow steps $i \to i+j$ with $j \in \{-1, 0, 1, \dots, k\}$ with weights $c_j \in \mathbb{R}$, then the adjacency matrices are of the form $\mathtt{toep}(0, \dots, 0, c_{-1}, c_0, \dots, c_k, 0, \dots, 0)$. For these, we can give linear recursions for the characteristic polynomials. With Fact 3 we obtain linear recursions for the number of corresponding lattice paths if we get a handle on the characteristic polynomial $\chi_n(x)$ of the $n \times n$ matrix $\mathtt{toep}(0, \dots, 0, c_{-1}, c_0, \dots, c_k, 0, \dots, 0)$.
A recursion for the characteristic polynomial can be obtained via Lemma 8. In a cycle decomposition the last vertex v_n belongs to a cycle $v_n \overset{-1}{\to} v_{n-1} \overset{-1}{\to} \dots \overset{-1}{\to} v_{n-j}, \overset{j}{\to} v_n$ for some j. This yields the recursion

$$\chi_{n+1}(x) = (c_0 - x)\chi_n(x) + \sum_{j=1}^{k} c_{-1}^j c_j \chi_{n-j}(x).$$

For small k we can solve the linear recursion above and get an explicit representation of the characteristic polynomial, and hence an explicit linear recursion for the numbers of lattice walks.

3.5.4 Cylindrical lattice walks

As a last instance we now consider walks, which allow arbitrary step length (and arbitrary weights) on a cycle. Taking an edge for each allowed step we obtain a circulant graph. The "time expansion" yields a cylinder as analog to the lattice strips.

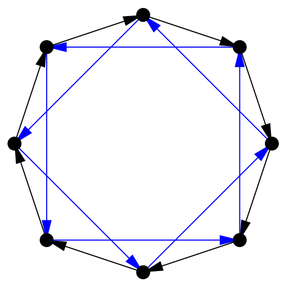

Figure 3.5: A circulant graph with 8 vertices and steps +1 (black) and −2 (blue).

The adjacency matrices of circulant graphs are *circulant matrices*. For this class of matrices eigenvalues and eigenvectors are known. According to Gray [27] they are as follows:

Theorem 13 *(Eigenvalues and -vectors of circulant matrices, Theorem 3.1 of [27])* *Let* $C \in \mathbb{R}^{n\times n}$ *a circulant matrix with first row* $(c_0, c_1, \dots, c_{n-1}) \in \mathbb{R}^n$*. The eigenvalues* ψ_m *and corresponding eigenvectors* y_m *of* C *for* $m = 0, 1, \dots, n-1$ *are as follows:*

$$\psi_m =: \sum_{k=0}^{n-1} c_k e^{\frac{-2\pi i m k}{n}}, \qquad y_m =: \frac{1}{\sqrt{n}}\left(1, e^{\frac{-2\pi i m}{n}}, e^{\frac{-2\pi i m 2}{n}}, \dots, e^{\frac{-2\pi i m(n-1)}{n}}\right).$$

Note, that the eigenvectors of circulant matrices do *not* depend on the matrix, but only on the dimension n. Therefore, they are the same for all circulant matrices of the same size (which implies these matrices commute). Now, we can count walks on circulant graphs with the techniques from Section 3.2. This leads to explicit formulas as well as to linear recursions.

3.6 Conclusion

Trigonometric sums of the form given in Theorem 10 can be handled quite well by computer algebra systems, e.g. `Maple`. Our impression is that indeed the formulas lead to the most effective way of evaluating the number of paths of a certain type in not too narrow and rather long strips. For narrow strips a generating function approach may be practical and superior. For short strips it can be reasonable to explicitly use the recursion (dynamic programming).
In the examples section (Section 3.3) we have listed some cases of integer sequences that could be obtained by counting (weighted) walks in $P_n(\alpha)$. The basic approach,

however, is not limited to this case. The crucial requirement is that we are able find an explicit expression for the entries of powers of the adjacency matrix. Cases where this is possible include situations where Toeplitz matrices are substituted for the entries of a Toeplitz matrix. Investigating this and related cases should allow to count families of *lattice surfaces*, e.g., fillings of the cells of an $n \times m$ grid with numbers between 0 and h such that the numbers of adjacent cells differ by at most one.

Some problems remain. For example in Section 3.4, we found a linear recursion, but the explicit generating functions remain unknown. In Section 3.5.1 we quote results for band matrices with (at most) five nonzero diagonals. Improvements in this work can be directly translated back into lattice path enumeration. Section 3.5.3 contains linear recursions for linear recursions, can this be simplified?

Another direction of improvement might be to try to solve special sets of bounded lattice walks, for example with a set of steps $i \to i + k$ with $k \in \{-3, -1, 0, +1, +5\}$. Banderier and Nicodeme considered in [5] some sets of this shape and were able to enumerate the corresponding number of lattice walks. The methods in this chapter do not cover these families of walks, but maybe it is possible, to understand the structure of the powers of the corresponding Toeplitz matrices.

The last direction of progress, we would like to mention, might be the study of the inversion of general Toeplitz matrices, since we can apply Theorem 12, which we already applied in the proof of Theorem 11.

In general the inverse of Toeplitz matrices are interesting. If they are given we can apply Theorem 12, similar to what we did in the proof of Theorem 11. There is a lot of work in this direction, see [16] and references therein.

4 Tower Moves

4.1 Preliminaries

Before diving into the results, this section reviews briefly some preliminaries which are specific to this chapter. First, recall the definition of α-orientations:

Definition 16 (α-orientations)
Let $G = (V, E)$ be a simple plane graph with vertex set V and edge set E. Further, let $\alpha : V \to \mathbb{N}$ be a weight function, living on the vertices of G. An *α-orientation* $\vec{G}$ of G is an orientation of the edge-set of G, such that every vertex $v \in V$ has out degree $\alpha(v)$. Denote the set of all α-orientations with $\Omega_{\alpha,G}$ or Ω for short, if G and α cannot be mistaken. We call α *feasible*, if there exists an α-orientation. Further, edges are called *rigid*, if they are oriented in the same way in all α-orientations $\Omega_{\alpha,G}$.

In addition, remember the bijection between α-orientations and α-potentials, which led to the notion of k–heights in Chapter 2:

Definition 17 (essential cycles and α-potentials)
More formally, a simple cycle C of G is an essential cycle if

- C is cord-free
- the interior cut of C is rigid
- there exists an α-orientation $\vec{X}$ of G, such that C is oriented in $\vec{X}$

Further, let $\mathcal{C}$ be the set of essential cycles and $\vec{G}_{\min}$ the minimum of Ω. In this case, $\phi : \mathcal{C} \to \mathbb{N}$ is an *α-potential* of G if

- $|\phi(C) - \phi(C')| \leq 1$ for all $C, C' \in \mathcal{C}$, which share an edge e
- $\phi(C) \leq 1$ if C is the only essential cycle containing an edge e
- if $C^{l(e)}$ and $C^{r(e)} \in \mathcal{C}$ are the cycles left and right of e in $\vec{X}_{\min}$, we have $\phi(C^{l(e)}) \leq \phi(C^{r(e)})$

For the graphs we consider, the essential cycles are the set of faces of the underlying graphs, except the outer face. We are now finally going to work on the face flip Markov chain living on the set of α-orientations of some fixed graph G. It is defined as follows:

Definition 18 (the face flip Markov chain)
Let G be a plane graph and α a weight-function for G, such that there is an α-orientation of G. Now the *face flip Markov chain M* operates as follows on Ω:

Input: $\vec{X} \in \Omega$
Output: $\vec{X} \in \Omega$
Let $p \longleftarrow [0,1]$ uniformly at random
Let $f \longleftarrow F'(G)$ a flippable face of G, uniformly at random
if f *is counter-clockwise oriented* ***and*** $p \leq \frac{1}{2}$ **then**
  reorient f clockwise, i.e. flip f
if f *is clockwise oriented* ***and*** $p > \frac{1}{2}$ **then**
  flip f
return $\vec{X}$

Algorithm 5: One step of the face flip Markov chain M.

This is a Markov chain on Ω. It is reversible, since a flipped face is again oriented. Therefore it is recurrent, since Ω is a distributive lattice, where all elements are connected by sequences face flips. Since it does not change the state with a probability of $\geq \frac{1}{2}$ at every state, it is aperiodic, which implies, it is ergodic, so it has a unique stationary distribution π. For the graphs, which we consider, the set Ω is a distributive lattice, which is connected by face flips, so M is irreducible and due to the symmetry of the chain, we see $\pi(\vec{X}) = \frac{1}{|\Omega|}$ for all $\vec{X} \in \Omega$. To be able to analyze the behaviour, especially the mixing time of M, we use a well known fact (for example, Chapter 7.2 of [38] covers this as *Bottleneck Ratio*), which in our words sounds as follows:

Definition 19 (hourglass-shaped sample spaces)
Let Ω be a sample space and M be a Markov chain on Ω. Partition Ω into Ω_r, Ω_g and Ω_b (r,g, and b stand for red, green and blue, which originates in the colour of an edge in a Schnyder wood of a graph). If Ω_g separates Ω_b from Ω_r, i.e. for all $s_1 \in \Omega_b$ and $s_2 \in \Omega_r$ the transition probability between s_1 and s_2 is 0 (in both directions), we say Ω is *hourglass-shaped* (with respect to the partition). This means every path of steps of M from Ω_r to Ω_b has to hit Ω_g at some time (as well as every path of steps from Ω_b to Ω_r).

Depicting Ω as the vertices of a graph and the transitions of M as edges, the resulting directed graph looks like an hourglass, if Ω_g is much smaller than Ω_b and Ω_r.
Let π be the stationary distribution of an symmetric and ergodic Markov chain M on Ω (i.e. the uniform distribution). Assuming $|\Omega_b| \gg |\Omega_g| \ll |\Omega_r|$ and $\pi(S) = \sum_{s \in S} \pi(s)$ for all $S \subseteq \Omega$ in an hoursglass-shaped Ω, we can easily give a bound on the *conductance*

$$\Phi_M := \min_{S \subseteq \Omega, \pi(S) \leq \frac{1}{2}} \frac{1}{\pi(S)} \sum_{s_1 \in S, s_2 \notin S} \pi(s_1) \cdot \Pr(s_1, s_2)$$

of M:

Lemma 9 *(conductance of hourglass shaped state spaces)*
Let M be an symmetric and ergodic Markov chain on an hourglass shaped state space $\Omega = \Omega_r \cup \Omega_g \cup \Omega_b$. We have:

$$\Phi_M \leq \frac{\pi(\Omega_g)}{\min\{\pi(\Omega_r), \pi(\Omega_b)\}}$$

Proof. Without loss of generality let $\pi(\Omega_r) \geq \pi(\Omega_b)$. This implies $\pi(\Omega_b) \leq \frac{1}{2}$ due to

$$2\pi(\Omega_b) \leq \pi(\Omega_r) + \pi(\Omega_b) \leq \pi(\Omega_r) + \pi(\Omega_g) + \pi(\Omega_b) = \pi(\Omega_r \cup \Omega_g \cup \Omega_b) = \pi(\Omega) = 1.$$

So Ω_b is a valid subset $S \subset \Omega$ with $\pi(S) \leq \frac{1}{2}$ as in the definition of Φ_M and we gain:

$$\begin{aligned}
\Phi_M &= \min_{S\subseteq\Omega, \pi(S)\leq\frac{1}{2}} \frac{1}{\pi(S)} \sum_{s_1\in S, s_2\notin S} \pi(s_1)\cdot \Pr(s_1,s_2) \\
&\leq \frac{1}{\pi(\Omega_b)} \sum_{s_1\in\Omega_b, s_2\in\Omega_r\cup\Omega_g} \pi(s_1)\cdot \Pr(s_1,s_2) \\
&= \frac{1}{\pi(\Omega_b)} \left(\sum_{s_1\in\Omega_b, s_2\in\Omega_r} \pi(s_1)\cdot \underbrace{\Pr(s_1,s_2)}_{=0} + \sum_{s_1\in\Omega_b, s_2\in\Omega_g} \pi(s_1)\cdot \underbrace{\Pr(s_1,s_2)}_{\leq 1} \right) \\
&= \frac{1}{\pi(\Omega_b)} \sum_{s_1\in\Omega_b, s_2\in\Omega_g} \frac{1}{|\Omega|}\cdot \Pr(s_1,s_2), \text{ due to ergodicity of } M \\
&= \frac{1}{\pi(\Omega_b)} \cdot \frac{1}{|\Omega|} \sum_{s_2\in\Omega_g} \underbrace{\left(\sum_{s_1\in\Omega_b,} \Pr(s_2,s_1) \right)}_{\leq 1}, \text{ due to symmetry of } M \\
&\leq \frac{1}{\pi(\Omega_b)} \cdot \frac{|\Omega_g|}{|\Omega|} \\
&= \frac{\pi(\Omega_g)}{\pi(\Omega_b)}, \text{ again due to ergodicity of } M \\
&= \frac{\pi(\Omega_g)}{\min\{\pi(\Omega_r), \pi(\Omega_b)\}}
\end{aligned}$$

□

Furthermore, there are well known relations between the conductance of a Markov chain M and its *mixing time*

$$\tau_M(\varepsilon) := \sup_{x\in\Omega} \min\{t : \|\mathcal{M}^t_{x,\cdot} - \pi\|_{TV} \leq \varepsilon\},$$

where $\mathcal{M}$ is the transition matrix of M as introduced in Chapter 1.
One well-known theorem linking these two, is the following:

Theorem 14 *(Jerrum and Sinclair; [34])*
Let M be a Markov chain on Ω with conductance Φ_M. The mixing time τ_M of M on Ω satisfies

$$\tau_M \geq \frac{1}{4\cdot\Phi_M} - \frac{1}{2}.$$

Note that the *Cesaro mixing time* $\tau_M := \tau_M(\frac{1}{4})$ does not depend on ε as in the previous definition, by setting it to some fixed value, [33], Remark 4.4 and [38] ,Theorem 6.15 justify this simplification, because we are only interested in the (exponential) grow rates of the bounds to be estimated. Another theorem, which will be applied later is a version of the coupling theorem of Dyer and Greenhill [17] by Bordewich and Dyer:

Theorem 15 *(Bordewich and Dyer; Theorem 7 with $\delta = 1$ and $S = \{(v,w) \in \Omega^2 \mid d(v,w) = 1\}$ of [8])*
Let P be a coupling for the Markov chain M and let $d : \Omega \times \Omega \to \mathbb{N}$ be a metric and $S = \{(v,w) \in \Omega^2 \mid d(v,w) = 1\}$. Let $\beta \leq 1$ such that

$$\mathbb{E}_P[d(X_{t+1}, Y_{t+1})] \leq \beta$$

for all $(X_t, Y_t) \in S$. Further, let D be the maximum value of d on $\Omega \times \Omega$ and

$$\widetilde{p} := \min_{(v,w)\in S} \Pr(X_{t+1} = w \mid X_t = v).$$

If $\widetilde{p} > 0$ the mixing time $\tau^\star(\varepsilon)$ of $M^\star = \frac{1}{1+\widetilde{p}} \cdot (M + \widetilde{p} \cdot \mathcal{I})$ fulfills

$$\tau_{M^\star}(\varepsilon) \leq \left\lceil \frac{eD^2(1+\widetilde{p})}{\widetilde{p}} \right\rceil \cdot \left\lceil \log(\varepsilon^{-1}) \right\rceil.$$

This theorem is a stronger version of the coupling theorem of Dyer and Greenhill (Theorem 2.1, [17]), which is applied by Fehrenbach, Rüschendorf [18] as well as Miracle, Randall, Streib and Tetali [39]. However Creed [12] already uses this improved version. Anyhow, as the following remark shows, the theorem of Dyer and Greenhill cannot easily be applied here (or even in the case of planar triangulations with maximum degree ≤ 6):

Remark 4
Figure 4.1 shows a 3-orientation of a planar triangulation with maximum degree 6, if you orient all blue edges clockwise (or all blue edges counter-clockwise). Furthermore, both graphs have the same set of oriented faces, i.e. a coupling of the face flip Markov chain would apply the same step to both orientations, which means, that in a coupling the probability of a distance-change is 0.
Therefore it cannot be bounded away from 0 and the Theorem of Dyer and Greenhill is not easily applicable. Still it might be applicable, if one considers also the tower-moves of the tower chain, which is going to be introduced later.

Theorem 15 will be applied in combination with the comparison theorem of Diaconis and Saloff-Coste [15] in the version of Randall and Tetali [42] to find a bound for the mixing time of M. We already saw this theorem in Chapter 2 as Theorem 6 on page 24. For ease, we shortly repeat it:

Theorem 16 *(Randall and Tetali, Proposition 4 of [42])*
Let M and $\widetilde{M}$ be two reversible Markov chains on the same state space Ω with the same stationary distributions π and $\pi_\star := \min_{x\in\Omega} \pi(x)$. Let $E(M)$ be the set of transitions of M and $E(\widetilde{M})$ be the set of transitions of $\widetilde{M}$. For each pair $(u,v) \in E(\widetilde{M})$ define a path γ_{uv}, which is a sequence $u = u_0, u_1, \ldots, u_k = v$ of transitions of M, i.e. $(u_i, u_{i+1}) \in E(M)$ for all i. For $(x,y) \in E(M)$ let

$$\Gamma(x,y) := \{(u,v) \in E(\widetilde{M}) \mid (x,y) \in \gamma_{uv}\}.$$

Further let

$$A := \max_{(x,y)\in E(M)} \left\{ \frac{1}{\pi(x)\cdot(\mathcal{M})_{(x,y)}} \cdot \sum_{(u,v)\in\Gamma(x,y)} |\gamma_{uv}| \cdot \pi(u) \cdot (\widetilde{\mathcal{M}})_{x,y} \right\}$$

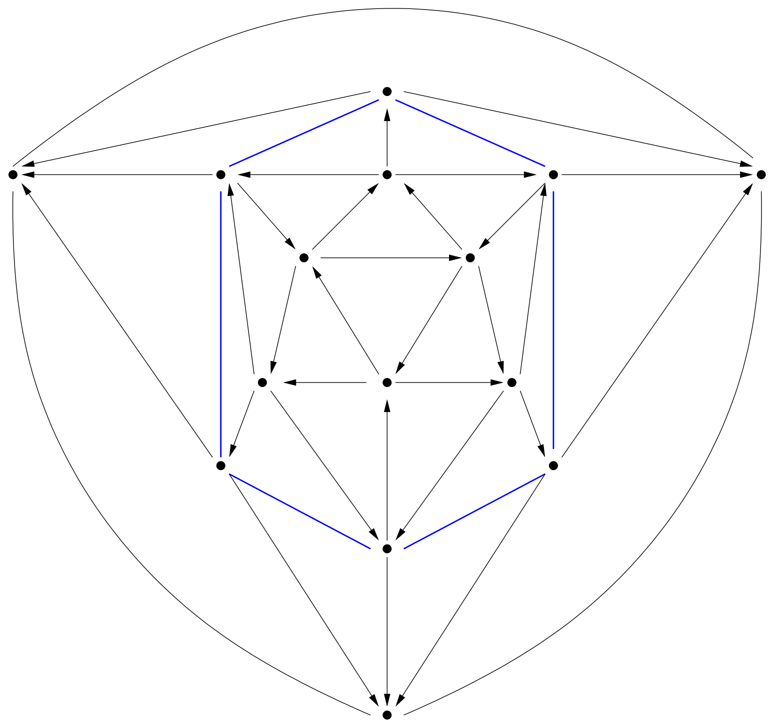

Figure 4.1: A plane triangulation with maximum degree 6, such that two different orientations only permit the same face flips.

where $(\mathcal{M})_{x,y} := \Pr_M(x_{t+1} = y \mid x_t = y)$*, i.e. the transition probability from* x *to* y *in one step / the entry associated to* x *and* y *in the transition matrix of* M*. If the second largest eigenvalue* λ_1 *of* $\widetilde{M}$ *complies with* $\lambda_1 \geq \frac{1}{2}$*, we know that for every* $0 < \varepsilon < 1$ *the mixing time* $\tau_M(\varepsilon)$ *is bounded by:*

$$\tau_M(\varepsilon) \leq \frac{4\log(\frac{1}{\varepsilon \cdot \pi_\star})}{\log(\frac{1}{2\varepsilon})} \cdot A \cdot \tau_{\widetilde{M}}(\varepsilon)$$

4.2 3-Orientations of planar triangulations

First, let us revisit the already mentioned, recent result of Miracle, Randall, Streib, and Tetali. We modify it slightly, by presenting not their family of graphs to prove the result, but a family, which –in our eyes– is more appealing. Also the family of graphs, which is applied, later on renders it possible to improve the result in terms of the maximum degree of the graphs' vertices. Their result covers 3-orientations of planar triangulations, which are α-orientations, such that $\alpha(v) = 3$ for all vertices $v \in V(G) \setminus \{a, b, c\}$, where a, b, and c are the vertices on the outer face of G. For them α is defined by $\alpha(a) = 0$, $\alpha(b) = 1$ and $\alpha(c) = 2$, so all three edges on the outer face are rigid as well as all other edges which contain a vertex on the outer face.

Theorem 17 *(Miracle et.al., [39])*
There is an infinite family of planar triangulations of graphs, such that the face flip

Markov chain does not rapidly mix the set of 3-orientations of these graphs, i.e.

$$\tau_M \geq 2^{(n-14)/4} - \frac{1}{2},$$

for a graph with n vertices.

In fact, we prove the following claim: There is an infinite family of planar triangulations $\{G_1, G_2, \dots\}$ such that the face flip Markov chain does not rapidly mix the set of 3-orientations of G_n, i.e.

$$\tau_M \geq 2^{n-3} - 1,$$

where G_n has $3n+4$ vertices for all $n \in \mathbb{N}$. To do so, we present a new family of graphs, which are simpler, than the ones, presented by Miracle et. al. and allow to decrease the maximal degree in a later corollary. Our proof sounds as follows:

Proof. Consider the graph G_n on $3n+4$ vertices, depicted in Figure 4.2.

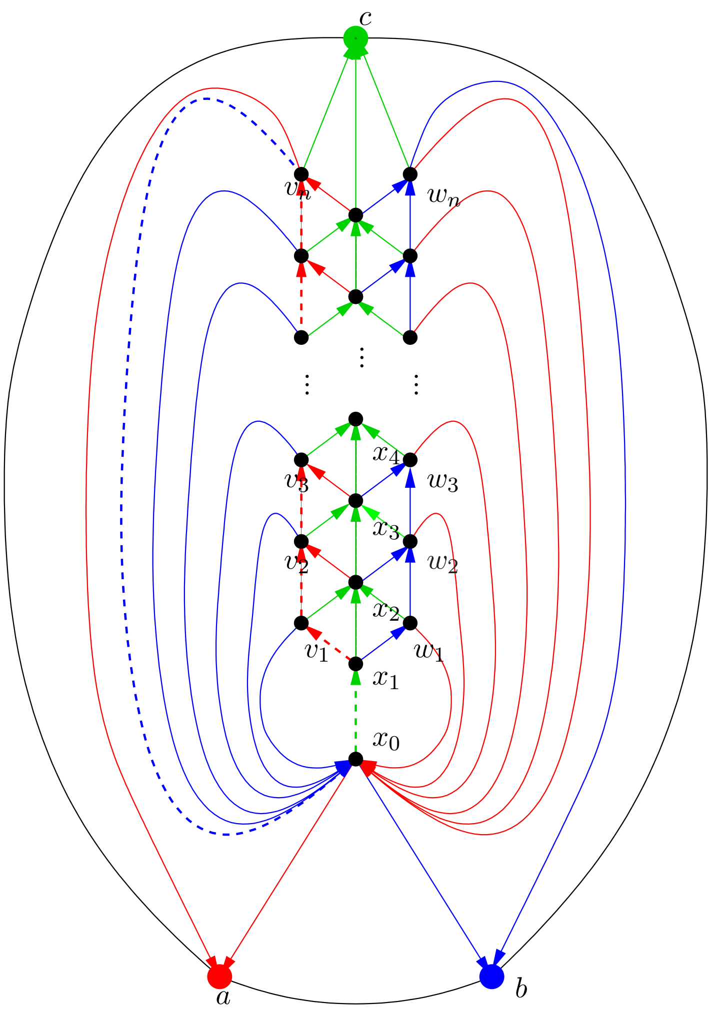

Figure 4.2: A 3-orientation / Schnyder-wood of G_n.

It is a planar triangulation and there are 3-orientations of G_n, since the picture shows an example. So $\alpha \equiv 3$ is feasible. Now, let Ω be the set of 3-orientations of G_n. To show that Ω is hour glass-shaped, consider the partition $\Omega = \Omega_g \cup \Omega_r \cup \Omega_b$ with respect to the colour of the edge $\{x_0, x_1\}$ in the Schnyder-wood, corresponding to the orientation, under the assumption that c is the green sink. This is equivalent to saying

$$\begin{aligned} \Omega_g &:= \{\vec{G} \in \Omega \mid (x_0, x_1) \in E(\vec{G})\}, \\ \Omega_b &:= \{\vec{G} \in \Omega \mid \exists i : (x_0, w_i) \in E(\vec{G})\}, \text{ and} \\ \Omega_r &:= \{\vec{G} \in \Omega \mid \exists i : (x_0, v_i) \in E(\vec{G})\}, \end{aligned}$$

since x_0 is adjacent to the red and to the blue sink, so these outgoing edges are rigid and the choice of the outgoing green edge of x_o fixes the colours of all other edges, adjacent to x_0.
Now, the argue for the claims $|\Omega_g| = 1$ and $|\Omega_r| = |\Omega_b| \geq 2^n$ as well as the one, that every path of face flips / Markov chain transitions from Ω_r to Ω_b (or vice versa) hits at least one element of Ω_g. Together this proves that Ω is an hour glass-shaped state set, so it has a bad conductance, which suffices to show that the face flip Markov chain on Ω is not rapidly mixing.

- Claim $|\Omega_g| = 1$: If $\{x_0, x_1\}$ is green, it has to be oriented towards this vertex. Since $\deg(x_1) = 4$ all other edges containing x_1 are directed away from x_1. Now the same holds for v_1 and w_1, so all edges between $\{x_1, v_1, w_1\}$ and $\{x_2, v_2, w_2\}$ have to be directed towards the second set (i.e. oriented upwards in the picture). Inductively, all edges between $S_i := \{x_i, v_i, w_i\}$ and S_{i+1} have to be directed towards S_{i+1}, if the edges from S_{i-1} are directed to S_i, because otherwise these three vertices of S_i can not reach their desired out-degree of 3. Finally, since x_0 already has out-degree 3, all edges $\{x_0, v_i\}$ and $\{x_0, w_j\}$ are directed towards x_0, which completes the picture. So $|\Omega_g| = 1$ and Figure 4.2 shows the only 3-orientation of G_n in Ω_g.

- Claim $|\Omega_r| = |\Omega_b| \geq 2^n$: Due to the symmetry with respect to the $x_0, x_1, \ldots, x_{n+1}, c$ - axis of G_n, every orientation can be mirrored on this axis. If we do this to a red or blue one, the outgoing green edge of x_0 is mapped from one side of the edge (x_1, x_0) to the other, so every element of Ω_r is mapped to an element of Ω_b and vice versa. This means, that this mirror-operation is a bijection between Ω_r and Ω_b, and they have the same cardinality.

  Furthermore, assume the directed cycle $x_0, x_1, v_1, v_2, \ldots, v_n, x_0$ in the element of Ω_g, which is dashed in the picture, is reversed. The resulting orientation is in Ω_r, where every triangle x_i, v_i, v_{i-1} can be flipped independently of each other and without leaving Ω_r. Therefore $|\Omega_r| \geq 2^{n-1}$. More detailed counting even yields $|\Omega_r| \geq 2^n$ easily.

- Claim: Ω_g separates Ω_b from Ω_r: Assume $\vec{G} \in \Omega_r$. To transform $\vec{G}$ into an element of Ω_b, we have to move the outgoing green edge of x_0 from some v_i to some w_j. Since (x_0, a) and (x_0, b) are rigid, this can only happen via $v_{i-1}, v_{i-2}, \ldots x_1, w_1, \ldots w_j$, since every face flip can only move it one step further. So on any path from an element of Ω_r to an element of Ω_b the outgoing green edge of x_0 has to be (x_0, x_1), which means, that this path hits Ω_g.

So Ω is an hour glass-shaped state lattice. Therefore Lemma 9 yields:

$$\Phi_M \leq \frac{\pi(\Omega_g)}{\pi(\Omega_b)} = \frac{|\Omega_g|}{|\Omega|} \cdot \frac{|\Omega|}{|\Omega_b|} \leq \frac{1}{2^{n-1}}$$

and Theorem 14 leads to

$$\begin{aligned} \tau_M &\geq \frac{1}{4 \cdot \Phi_M} - \frac{1}{2} \\ &\geq \frac{2^{n-1}}{4} - \frac{1}{2} \\ &\geq 2^{n-3} - 1. \end{aligned}$$

□

Note, that G_n has one single vertex of degree $2n+3$, namely x_0, while all other vertices have a degree of at most 6. Since the face flip Markov chain is rapidly mixing on all 3-orientation lattices of planar triangulation with maximum degree 6, this is rather close. Furthermore, the next corollary shows that the maximum degree can even be pressed down the size to $O(\sqrt{n})$:

Corollary 5
There is an infinite family of planar triangulations $\mathcal{F}' = \{G'_n \mid n \in \mathbb{N}, n \geq 2\}$ such that G'_n has $2n^2 + n + 4$ vertices, a maximum degree of $2n + 4$ and a mixing time

$$\tau_M \geq 2^{(n-1)\log(2n+2)-2} - \tfrac{1}{2},$$

which means, the face flip Markov chain M is not rapidly mixing any of the sets of 3-orientations of a graph in this family.

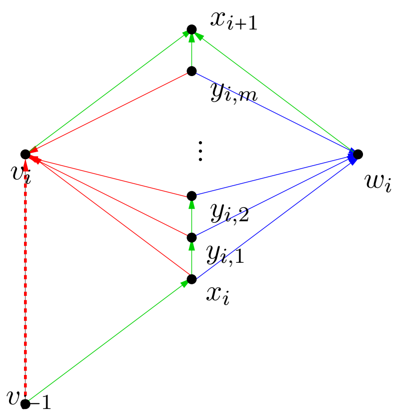

Figure 4.3: The subdivision of $\{x_i, x_{i+1}\}$.

Proof. Starting with the graphs G_n already introduced in Theorem 17, the edges $\{x_i, x_{i+1}\}$ are subdivided for all i as shown in Figure 4.3, i.e. for all $i = 1, \ldots, (n-1)$ we introduce m vertices $y_{i,j}$ for $j = 1, \ldots, m$ and connect x_i and $y_{i,1}$ as well as x_{i+1} and $y_{i,m}$. Furthermore $y_{i,j}$ is connected to $y_{i,j+1}$ (for $j \neq m$), v_i and w_i. The resulting graph is called $G_{n,m}$. This graph has $3n + 4 + (n-1) \cdot m$ vertices. The maximum degree is $\max\{2n+3, m+4\}$ and it has one vertex of degree $2n+3$ and $2n$ vertices of degree $m+4$. We claim, that the set of 3-orientations of $G_{n,m}$ is still hour glass-shaped. The partition is the same already applied in the previous proof.

- Claim $|\Omega_g| = 1$: If $\{x_0, x_1\}$ is green, it has to be oriented towards x_1. This vertex has degree 4, so all other adjacent edges are directed away from x_1. Now, the same is true for $y_{1,1}$ because $(x_1, y_{1,1})$ is directed towards $y_{1,1,}$ and it has a degree of four. Inductively it holds for $y_{1,i}$ up to $y1, m$. So all edges are oriented upwards, as indicated in the picture. Now, due to the degree of v_1 and w_1, the edges (v_1, x_2) and (w_1, x_2) are oriented towards x_2. This means, the edge from x_1 to x_2 is again directed upwards, which inductively forces all edges to be oriented as indicated in Figure 4.2 with Figure 4.3 inserted $n-1$ times. This finally yields $|\Omega_g| = 1$.

- Claim $|\Omega_r| = |\Omega_b| \geq (m+2)^{n-1}$: Since the inserted tokens do not break symmetry, the bijection between Ω_r and Ω_b which was made use of in Theorem 17 still exists. Therefore both sets still have the same size.

  Further, the directed cycle $x_0, x_1, v_1, v_2, \dots, v_n, x_0$ can be reversed. This makes all edges $\{v_i, v_{i+1}\}$ blue. Now there are m+1 directed cycles $c_k = \{x_i, y_{i,1}, \dots, y_{i,k}, v_i, v_{i-1}\}$, which we can flip one of. Additionally we can also flip the triangle $\{v_i, v_{i-1}, x_i\}$, which gives us $m+2$ choices for every $i \in \{1, \dots, n-1\}$, independent of each other, explaining $|\Omega_r| \geq (m+2)^{n-1}$.

- Claim: Ω_g separates Ω_b from Ω_r: Since $G_{n,m}$ has the same local structure around x_0 as G_n and there Ω_{g,G_n} separates Ω_{r,G_n} from Ω_{b,G_n}, the same is true here.

This means, Lemma 9 is applicable and combined with and Theorem 14 it provides:

$$\begin{aligned}\tau_M &\geq \frac{1}{4 \cdot \Phi_M} - \frac{1}{2} \\ &\geq \frac{\min\{|\Omega_b|, |\Omega_r|\}}{4 \cdot |\Omega_g|} - \frac{1}{2} \\ &= \frac{(m+2)^{n-1}}{4} - \frac{1}{2}\end{aligned}$$

Now choosing $m = 2n$ and defining $G'_n = G_{2n,n}$ directly yields the claimed result. □

This finishes our section on 3-orientations of plane triangulations. The interesting open question is obviously for which maximum degrees the face flip Markov chain is mixing and for which it is not, since the gap between 6 and $\mathcal{O}(\sqrt{n})$ is still wide. This question is left open and we turn towards 2-orientations of plane quadrangulations to see the analogous picture there.

4.3 2-Orientations of plane quadrangulations

This section contains two results on the behaviour of the face flip Markov chain on the set of 2-orientations of a plane quadrangulation. These are embedded planar graphs, where all faces are 4-gons, whereas a 2-orientation is an α-orientation with $\alpha(v) = 2 \forall v \in V \setminus \{b, d\}$, where b and d are opposing vertices on the outer face. They have an out-degree of 0. This family of graphs is a generalisation of sub-sections of the rectangular grid-graph, where the Markov chain's behaviour was already investigated by Fehrenbach and Rüschendorf [18] in 2004. Even earlier the set of 2-orientations was considered with

analytic methods in the theory of Ising-models, for example in chapter 7 of Baxter's book [6] in 1982.
The first result shows, that for an unrestricted maximum degree we cannot expect the Markov chain to be rapidly mixing.
Afterwards, the second result proves, that the Markov chain is rapidly mixing the set of 2-orientations, if the graph has a maximum degree of 4. One can argue that this family is really small, but still it arouse some interest, even beside the special case of Ising-models. For example Hasheminezhad and Mc Ray [29] in 2010 published an algorithm to generate all these graphs. We start with some facts about plane quadrangulations:

Fact 4
A plane quadrangulation with n vertices has $2n-4$ edges and $n-2$ faces. Furthermore if G is 3-connected, it has $n-7$ essential cycles, i.e. faces, which can be flipped. Namely, these are all faces, except the outer face and the four faces, which share an edge with the outer face. Let $F'(G)$ be this set of flippable faces.

Also, note that 2-orientations of plane quadrangulations yield an edge-bicolouring analogue to Schnyder-Woods. An overview on these separating decompositions can be found in Felsner, Kappes, Huemer, and Orden [24]. According to [21] the study of separating decompositions goes back to [44] and continues in [43], [14], [13] and [22].
So let us start with the analog of Theorem 17, covering the face flip Markov chains behaviour for high maximum-degree plane quadrangulations:

Theorem 18
There is an infinite family of planar quadrangulations $\{Q_1, Q_2, \dots\}$ such that the face flip Markov chain does not rapidly mix the set of 2-orientations of any Q_n, i.e.

$$\tau_M \geq 2^{n-3} - 1,$$

while Q_n has $5n+5$ vertices.

Proof. This proof follows the line of arguments already applied in Theorem 17. So first, consider the graph Q_n in Figure 4.4. It has $5n+5$ vertices and is obviously a plane quadrangulations. All vertices except x_0 have a degree ≤ 4, while $\deg(x_0) = 2n+2$. Further, $\{a, x_0\} \in E$ and a is on the outer face, so x_0 has one rigid outgoing edge and another one, which is *free*. Again, we partition Ω, the set of 2-orientations of Q_n, into Ω_g, Ω_r and Ω_b, to show, that Ω is hour glass-shaped with respect to this partition. We partition as follows:

$$\begin{aligned}
\Omega_g &:= \{\vec{X} \in \Omega \mid (x_0, x_1) \in E(\vec{X})\}, \\
\Omega_r &:= \{\vec{X} \in \Omega \mid \exists i : (x_0, v_i) \in E(\vec{X})\}, \text{ and} \\
\Omega_b &:= \{\vec{X} \in \Omega \mid \exists i : (x_0, w_i) \in E(\vec{X})\}.
\end{aligned}$$

This means Ω_g contains of all orientations, such that the free outgoing edge of x_0 goes to x_1. Further Ω_r is the set of all orientations, where the free outgoing edge of x_0 is *left* of $\{x_0, x_1\}$, and Ω_b is the set, where the free edge is in the right hand side. Obviously, this is a decomposition of Ω into three disjoint sets. So the goal is again to show, that Ω is hour glass shaped:

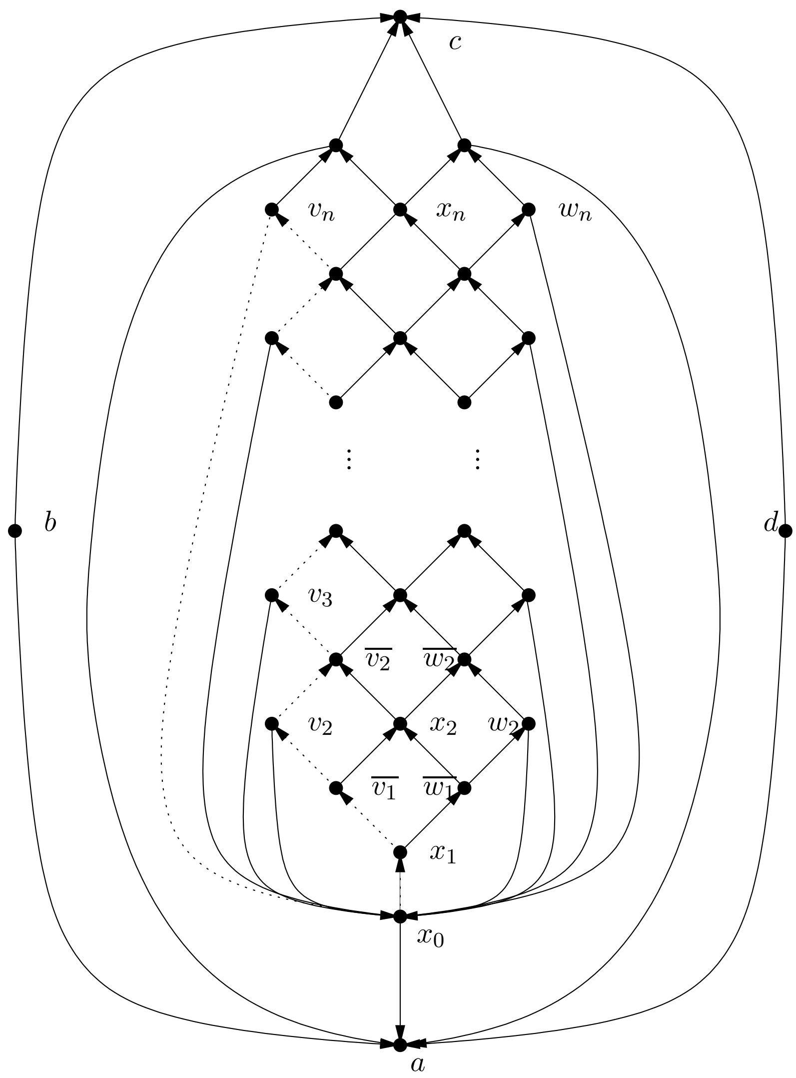

Figure 4.4: A 'green' 2-orientation of the graph Q_n.

- Claim $|\Omega_g| = 1$: If $\{x_0, x_1\}$ is green, it has to be oriented towards x_1. Now, we work from top to bottom. The vertices $\overline{v_{n+1}}$ and $\overline{w_{n+1}}$ are adjacent to c and a, so they both have two rigid outgoing edges. Now the edges from $\{v_n, x_n, w_n\}$ to $\{\overline{v_{n+1}}, \overline{w_{n+1}}, x_0\}$ have all to be directed towards the second set, i.e. upwards. Inductively all edges from $\{\overline{v_i}, \overline{w_i}\}$ to $\{v_{i+1}, x_{i+1}, w_{i+1}\}$ are directed upwards as well as all edges from $\{v_i, x_i, w_i\}$ to $\{\overline{v_i}, \overline{w_i}, x_0\}$, which leaves us no choice at all, except the one presented in Figure 4.4.

- Claim $|\Omega_r| = |\Omega_b| \geq 2^n$: Due to the symmetry with respect to the $a, x_0, x_1, \ldots, x_n, c$ - axis of G_n, the graph can be mirrored on this axis. We can also mirror any orientation on this axis, since no edge crosses it. Doing so, the free outgoing edge of x_0 is mapped from one side of the edge (x_1, x_0) to the other. This yields a bijection, mapping every element of Ω_r to an element of Ω_b and vice versa, and so both sets have the same cardinality.

Next, reversing the dashed directed cycle $x_0, x_1, \overline{v_1}, v_1, \overline{v_2}, v_2, \ldots, v_n, x_0$ in the unique element of Ω_g is possible. Doing so, we get an orientation in Ω_r, since the free outgoing edge of x_0 now ends in v_n. Now every face $\overline{v_i - 1}, x_i, \overline{v_i}, v_i$ is directed. So they can be flipped independently of each other without leaving Ω_r. Therefore $|\Omega_r| \geq 2^{n-1}$. More detailed counting even yields $|\Omega_r| \geq 2^n$ easily.

- Claim: Ω_g separates Ω_b from Ω_r: To transform $\vec{X} \in \Omega_r$ into an element of Ω_b, we have to move the free outgoing edge of x_0 from some v_i to some w_j. Since (x_0, a) is rigid, this can only happen via $v_{i-1}, v_{i-2}, \ldots x_1, w_1, \ldots w_j$, since every face flip can only move it one vertex further. So on any path from an element of Ω_r to an element of Ω_b the free outgoing edge of x_0 has to be (x_0, x_1), which means, that every path from Ω_r to Ω_b hits Ω_g at some step.

So Ω is an hour glass-shaped state lattice. Therefore, as in Theorem 17, the application of Lemma 9 and Theorem 14 yields

$$\begin{aligned}
\tau_M &\geq \frac{1}{4 \cdot \Phi_M} - \frac{1}{2} \\
&\geq \frac{\min\{|\Omega_b|, |\Omega_r|\}}{4 \cdot |\Omega_g|} - \frac{1}{2} \\
&= \frac{2^{n-1}}{4} - \frac{1}{2} \\
&\geq 2^{n-3} - 1,
\end{aligned}$$

which finishes this proof. □

After we have seen, that for plane quadrangulations with high maximum degrees the face flip Markov chain might be not rapidly mixing, we will see next the theorem, which states, that for low degrees it is rapidly mixing:

Theorem 19
Let G be a plane quadrangulation with n vertices, maximum degree ≤ 4 (except the vertices of the outer face), such that all essential cycles are facial cycles. The face flip chain is rapidly mixing on the set of 2-orientations of G, i.e.

$$\tau_M(\varepsilon) \leq \frac{4e(2n + \log(\varepsilon^{-1}))}{\log(\frac{1}{2\varepsilon})} \cdot n^6 \cdot \log(\varepsilon^{-1}).$$

The proof of this theorem follows the the same line of arguments as already applied in [12] and [39] for 3-orientations of planar triangulations and in [18] for 2-orientations of subsections of the rectangular grid. A short outline of the proof looks as follows:

$$\begin{array}{ccccccc}
\text{tower moves} & & \text{coupling} & & \text{Bordewich \& Dyer} & & \text{Randall \& Tetali} \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
M \longrightarrow & M_T & \longrightarrow & P & \longrightarrow & \tau_{M_T^*}(\varepsilon) & \longrightarrow \tau_M(\varepsilon)
\end{array}$$

Starting with the face flip Markov chain M, a tower Markov chain M_T, is introduced.

It can be coupled successfully. Next, using the coupling, a bound on the mixing time $\tau_{M_T^\star}(\varepsilon)$ is found with the theorem of Bordewich and Dyer. Here $M_T^\star$ is a further modification of M_T, according to Theorem 15. Now a comparison argument of Randall and Tetali finally allows to finish the proof.
Since the condition on the minimum degree of the graphs is not an issue, assume from now on that G is always is simple plane quadrangulation with maximum degree ≤ 4 and minimum degree ≥ 3. We continue by introducing some machinery, which allows us to prove the theorem:

Definition 20 (states of a face)
Each face of $\vec{G}$ in a 2-orientation is bounded by four oriented edges. These assign a *state* depending on the number of clockwise and counterclockwise edges to the face:

- the face is *oriented* if all edges are clockwise or all edges are counterclockwise oriented
- the face is *blocked* if all but one edge are clockwise oriented or all but one edge are counterclockwise oriented The edge oriented in the opposing direction is called the *blocking edge* of the face
- the face is *scrambled* if two edges are clockwise and two are counterclockwise oriented

Definition 21 (towers)
Let $\vec{G}$ be a 2-orientation of a plane quadrangulation G. A *tower* T is a path $(a_k, \dots, a_0)$ of faces of $\vec{G}$, such that

- a_0 is oriented
- $a_1, a_2, \dots a_k$ are blocked
- for all $i = 1, \dots, k-1$ the edge of a_i, opposing a_i's blocking edge, is the blocking edge of a_{i+1} (and a_2's blocking edge is an edge of a_0)

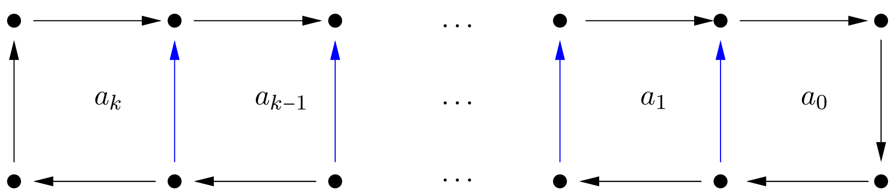

Figure 4.5: A tower with blue blocking edges.

The *length* of a tower *starting* in a_k and *ending* in a_0 is k. Finally, a tower is oriented (counter-) clockwise if the a_0 is oriented (counter-) clockwise.

This definition implies that the edges of a tower form one directed cycle plus inner blocking edges. Also a tower has a fixed direction to follow, after the starting face is fixed. Further, no two maximal towers can have a face –except a_0– in common, so two towers are either disjoint, share their oriented face, or one tower is a subset of the other.

Lemma 10 *(towers do not touch themselves)*
Let $T = (a_k, \ldots, a_0)$ be a tower in $\vec{G}$. Every edge e, which is contained in more than one face of T is a blocking edge of T.

Proof. By definition, the blocking edge of a_i is contained in a_i and a_{i-1}, so every blocking edge is contained in at least two faces of T. On the other hand, every face can only be contained once, so both faces being adjacent to the blocking edge are already part of T. Therefore the blocking edge cannot occur a third time in the tower, meaning: Blocking edges are contained exactly twice.

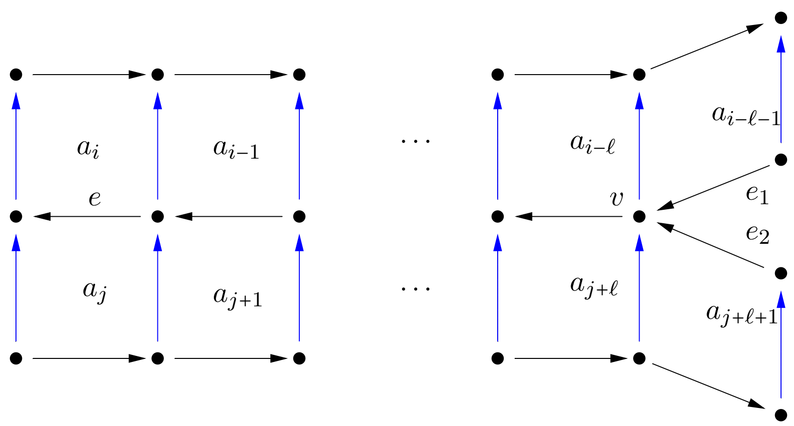

Figure 4.6: The edge e is contained twice and not blocking $\rightsquigarrow$ the vertex v has degree ≥ 5.

Assume e is an edge of a_i and a_j (with $i > j$) and not a blocking edge of T. So either a_{i-1} and a_{j+1} also share a non-blocking edge, or not. Since the blocking edges of a_m and a_{m+1} are opposing each other, they are disjoint. This means, due to finiteness, there must be some ℓ, such that $a_{i-\ell}$ and $a_{j+\ell}$ share a non-blocking edge and $a_{i-\ell-1}$ and $a_{j+\ell+1}$ do not. Let v be the starting vertex of the edge shared by $a_{i+\ell}$ and $a_{j-\ell}$. This implies, v is adjacent to this shared edge, the blocking edges of $a_{i-\ell}$ and $a_{j+\ell+1}$ and to the edges e_1 and e_2 which border a_{i+1} and a_{j-1}, so w has to have at least a degree of 5, which cannot be, since we are considering 2-orientations of plane quadrangulations of maximum degree 4. This situation is shown in Figure 4.6. □

This proof in fact could be reorganised, such that the maximum degree condition is not needed. It would contain a case distinction, taking into consideration the orientation of the blocking edges and the direction, in which the towers is moving. This reveals, that in any case moving direction and orientation of blocking edges does not fit together. Still, the next lemma makes use of the maximum degree condition, so only the shorter proof of Lemma 10 is presented here.

Corollary 6
A tower $T = (a_k, \ldots, a_0)$ can be flipped. First flip a_0, than a_1 ... and finally a_k. This results in a tower $T' = (a_0, a_2, \ldots, a_k)$.

Proof. Due to Lemma 10 only blocking-edges are covered twice by T, so if a_0 up to a_j are flipped, only the orientation of the blocking edge of a_{j+1} and part of the edges of the outer cycle of T have changed, while the other three edges of a_{j+1} remain untouched, meaning, a_{j+1} is now oriented and can be flipped. □

Lemma 11
Let $T = (a_k, a_{k-1}, \ldots, a_0)$ be a tower in $\vec{G}$. There are at most four oriented faces touching T, three of them at a_k and one at a_0.

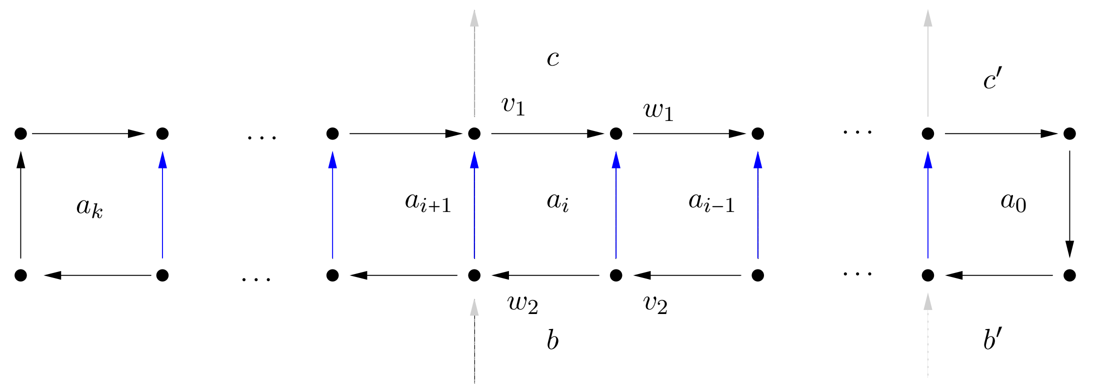

Figure 4.7: The situation at a_i (and a_0).

Proof. Consider a face $a_i, i \notin \{0, k\}$. The face has four neighbouring faces, two of them are a_{i-1} and a_{i+1} and two, which will be called b and c. Let $e_1 = (v_1, w_1)$ and $e_2 = (v_2, w_2)$ be edges of a_i, which are shared with b and c as shown in Figure 4.7.
Now, w_2 has already an out-degree of 2, so it might have another edge coming in, but for sure no further outgoing edge. If it has another incoming edge, the face b can therefore not be oriented, since the incoming edge is oriented against (v_2, w_2). If it does not have an incoming edge, b is also a neighbour of a_{i+1}. If $a_{i+1} = a_k$. This implies, b is a neighbour of a_k and is allowed to be oriented. Otherwise, by induction, either b is a neighbour of a_k or we have an incoming edge preventing b from being oriented.
On the other hand, v_1 has out-degree one. Therefore, it must have another outgoing edge, which is oriented against (v_1, w_2). So c is not oriented neither. This is depicted in Figure 4.7.
Next, consider the face a_0. Two neighbouring faces (b' and c') cannot be oriented, using the same argument as above, while the one opposing to a_1's blocking edge might be oriented. This finishes the proof. □

Note, that this lemma fails completely if the maximum degree was not bounded by 4! Since it is crucial to prove the theorem, it shows, that this machinery cannot easily be extend to plane quadrangulations with a maximum degree of 5 (or higher). Summing up the previous lemmata, it can be said, that towers behave *nicely*. So it is time set up our tower Markov chain M_T, which is an extension of the face flip Markov chain M:

Definition 22 (the tower Markov chain M_T)
One step of the tower Markov chain M_T looks as follows:

Input: $\vec{X} \in \Omega$
Output: $\vec{X} \in \Omega$
Let $p \longleftarrow [0,1]$ uniformly at random
Let $f \longleftarrow F'(G)$ a flippable face of G, uniformly at random
if *f is counter-clockwise oriented* ***and*** $p \leq \frac{1}{2}$ **then**
  flip f
if *f is clockwise oriented* ***and*** $p > \frac{1}{2}$ **then**
  flip f
if *a counter-clockwise oriented tower $T = (f, a_{k-1}, \ldots, a_0)$ starts in f* ***and*** $p \leq \frac{1}{4(k+1)}$ **then**
  reorient the outer circle of T clockwise, i.e. *flip* T
if *a clockwise oriented tower $T = (f, a_{k-1}, \ldots, a_0)$ starts in f* ***and*** $p > 1 - \frac{1}{4(k+1)}$ **then**
  flip T
return $\vec{X}$

Algorithm 6: One step of the tower Markov chain M_T.

The set of transitions of M_T is a superset of the set of transitions of M, so M_T is also irreducible. The chain is also aperiodic for the same reason and therefore it is ergodic. Further, a flipped tower is again a tower of the same length, so M_T is reversible, yielding $\pi(\vec{X}) = \frac{1}{|\Omega|}$ for all $\vec{X} \in \Omega$. Furthermore its second biggest Eigenvalue λ_1 is at least $\frac{1}{2}$, because independently from the choice of f and p in $\frac{1}{2}$ of all cases M_T does nothing. Now, to find a bound for the mixing time of M_T, consider the following path coupling of M_T:

Definition 23 (the path coupling of M_T)
Let P be our coupling, i.e. a Markov chain on $\Omega \times \Omega$, with transitions as follows: With $(\vec{X}, \vec{Y}) \in \Omega^2$ one step of P looks as follows:

Input: $(\vec{X}, \vec{Y}) \in \Omega \times \Omega$
Output: $(\vec{X}, \vec{Y}) \in \Omega \times \Omega$
Let $p \longleftarrow [0, 1]$ uniformly at random
Let $f \longleftarrow F'(G)$ a flippable face of G, uniformly at random
foreach $\vec{Z} \in (\vec{X}, \vec{Y})$ **do**
  if *f is ccw oriented in $\vec{Z}$* ***and*** $p \leq \frac{1}{2}$ **then**
    flip f in $\vec{Z}$
  if *f is cw oriented in $\vec{Z}$* ***and*** $p > \frac{1}{2}$ **then**
    flip f in $\vec{Z}$
  if *a ccw oriented tower $T = (f, a_{k-1}, \dots, a_0)$ starts in f in $\vec{Z}$* ***and*** $p \leq \frac{1}{4(k+1)}$ **then**
    flip T in $\vec{Z}$
  if *a cw oriented tower $T = (f, a_{k-1}, \dots, a_0)$ starts in f in $\vec{Z}$* ***and*** $p > 1 - \frac{1}{4(k+1)}$ **then**
    flip T in $\vec{Z}$
return $(\vec{X}, \vec{Y})$

Algorithm 7: One step of the coupling P of the tower Markov chain M_T.

Here, *cw* abbreviates clockwise and *ccw* does the same with counter clockwise. This Markov chain on $\Omega \times \Omega$ is indeed a (markovian) coupling of M_T, because it operates on $\vec{X}$ (and $\vec{Y}$) exactly in the same manner as Algorithm 6 does.
This means, P can be used to show that M_T is rapidly mixing, by applying Theorem 15. This theorem relies on a metric d on Ω, which in this case is the length of the shortest path between two elements in Ω, where we interpret Ω as a graph or distributive lattice. So $d(\vec{X}, \vec{Y})$ is the minimal number of face flips, we have to perform to transfer $\vec{X}$ into $\vec{Y}$, which is also the minimal number of steps of M we have to apply to $\vec{X}$ to achieve the same. So the set S of pairs of orientations at distance 1 is exactly the set of transition steps of the face flip Markov chain M, i.e. pairs of orientations, which differ in the orientation of one directed face. Let S be the set of these pairs at distance 1. Besides the metric d, we have to bound the parameters β, p and D of Theorem 15 for the tower chain M_T, as specified above.

Lemma 12 *(a bound for the maximum distance D in Ω)*
Let D be the maximum distance in Ω with respect to d.

$$D \leq \frac{n^2 - 4n + 4}{2}$$

holds.

Proof. To bound D, the maximum distance between two orientations, the bijection between α-orientations of planar graphs and their potentials is crucial. Further the fact, that all orientations form a distributive lattice, pointed out by Felsner in [20] is used and the fact, that every element of a distributive lattice is element of a shortest path from the minimum to the maximum of the lattice.
First, note that between any two orientations there is a path of transitions, which is not longer than the path from the lattice's minimum to its maximum. To prove this,

let h be the height of the lattice, i.e. the length of the shortest path from minimum to maximum. Any two orientations X and Y have a shortest path to the minimum of length a_X and a_Y and one to the maximum of length b_X and b_Y. Now, the distributivity of the lattice yields $h = a_X + b_X = a_Y + b_Y$, so $(a_X + a_Y) + (b_X + b_Y) = 2h$, meaning, that either the path from X to Y first using the shortest path from X to the minimum and followed by the reversed path from Y to it, or the corresponding path via the maximum has to have a length of at most h. So it suffices to bound the distance from minimum to maximum.
To do so, the bijection from α-orientations to α-potentials is applied. The minimum of the lattice corresponds to the 'all 0' potential. On the other hand, the potential of a face is at most its distance to the outer face, which is at most the radius of the dual graph. But the radius of G's dual is for sure not bigger than $\frac{1}{2}$ of the number of faces. So the α-potential corresponding to the maximum of the lattice has face entries of at most $\frac{\#\text{faces}}{2}$. This leads to

$$D \leq \underbrace{n-2}_{\#\text{faces}} \cdot \underbrace{\tfrac{1}{2}(n-2)}_{\geq\text{max potential value}} = \tfrac{n^2-4n+4}{2}.$$

□

The next example shows, that indeed there are plane quadrangulations, such that the distance is quadratic in in the number of vertices, as the previous lemma indicated.

Remark 5
Consider the family of graphs pictured in Figure 4.8. It is a plane quadrangulation with

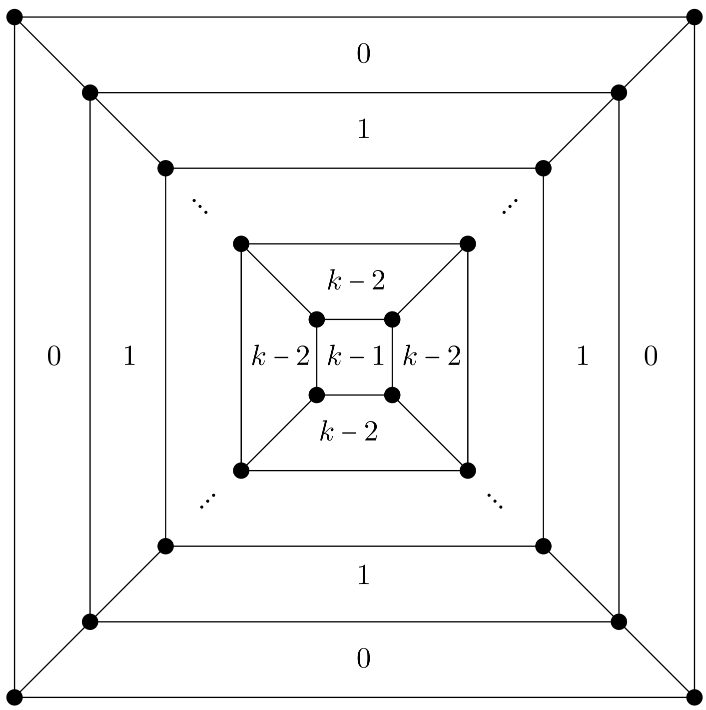

Figure 4.8: A family of graphs with large maximum distance D in Ω.

$n = 4k$ vertices and therefore $4k-2$ faces. The diameter is obviously k. As the numbers in the faces indicate, we have a potential, whose values sum up to

$$4 \cdot \sum_{j=0}^{k-2} j + k = 4\tfrac{(k-2)(k-1)}{2} + k = 2k^2 - 4k = \tfrac{n^2-2n}{8}.$$

So, for this family of graphs, the bound for D is tight up to a constant factor of 4 and therefore the bound of Lemma 12 can at most be improved for all graphs by a constant factor.

Lemma 13 *(estimating $\widetilde{p}$)*

$$\widetilde{p} := \min_{(\vec{X},\vec{Y})\in S} \Pr(X_{t+1} = \vec{Y} \mid X_t = \vec{X}) = \frac{1}{2(n-7)}.$$

Proof. The probability to move from $\vec{X}$ to $\vec{Y}$ in one step is exactly the probability to choose the directed face δ in which $\vec{X}$ and $\vec{Y}$ differ as well as the orientation, this face has in $\vec{Y}$. So the probability is $\frac{1}{2(n-7)}$, because we have $n-7$ flippable faces in $F'(G)$ and two choices for the orientation. □

Lemma 14 *(bounding β)*
Let $(\vec{X},\vec{Y}) \in S$, i.e. a pair of orientations, which differ only in the orientation of one oriented face. Then

$$\mathbb{E}_P[d(X_{t+1}, Y_{t+1}) \mid (\vec{X},\vec{Y}) = (X_t, Y_t)] \leq 1,$$

and with $\beta = 1$ we have

$$\mathbb{E}_P[d(X_{t+1}, Y_{t+1}) \mid (\vec{X},\vec{Y}) = (X_t, Y_t)] \leq \beta \cdot d(X_t, Y_t).$$

Proof. Let $\Delta := \mathbb{E}_P[1 - d(X_{t+1}, Y_{t+1}) \mid (\vec{X},\vec{Y}) = (X_t, Y_t)]$ be the expected change of distance between $\vec{X}$ and $\vec{Y}$ after one step of the coupling. Remember, $(\vec{X},\vec{Y}) \in S$ implies, they are at distance 1. This means, they differ only in the orientation of one oriented face $\delta \in F'(G)$. We consider different cases, depending on the relation of the oriented face δ to the face f, which the coupling chooses, to determine Δ for all possible configurations:

1. If $f = \delta$, the chosen face is oriented in $\vec{X}$ and in $\vec{Y}$. It is oriented clockwise in one instance and counter clockwise in the other one. Therefore, depending on the value of p, the face δ is flipped either in $\vec{X}$ or in $\vec{Y}$, while it is not flipped in the other. So in any case, the coupling's step result in two identical orientations, which results in
$$\Delta_1 = -1.$$
2. If $f \neq \delta$ and δ and f share an edge, some different situations can occur:
   a) If f is oriented in $\vec{X}$ or in $\vec{Y}$, assume without loss of generality it is cw oriented in $\vec{X}$. So in $\vec{Y}$ there starts a cw tower in f of length 1, which ends in δ, as depicted in Figure 4.9. So, if $p \leq \frac{1}{2}$, neither $\vec{X}$ nor $\vec{Y}$ nor their distances changes. For $\frac{1}{2} < p \leq \frac{7}{8}$ only f in $\vec{X}$ is flipped, which increases the

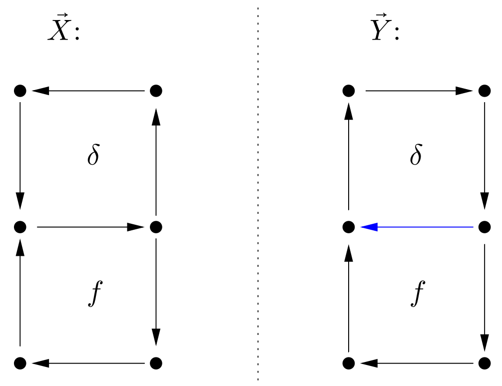

Figure 4.9: Situation 2.a, f is oriented in $\vec{X}$.

distance by 1. Finally, if $p > \frac{7}{8}$, we flip f in $\vec{X}$ and the tower (f, δ) is flipped in $\vec{Y}$, which decreases the distances by 1. This leads to

$$\Delta_{2a} = \tfrac{1}{2} \cdot 0 + \left(\tfrac{7}{8} - \tfrac{1}{2}\right)(+1) + \tfrac{1}{8}(-1) = \tfrac{1}{4}.$$

b) If f is scrambled in $\vec{X}$ (or $\vec{Y}$), f is blocked in $\vec{Y}$ and it might be, that a tower of length k starts in the orientation $\vec{Y}$ in f. So, no matter what p is, $\vec{X}$ remains unchanged, while there might be a tower getting flipped in $\vec{Y}$ which would increase the distance by $k+1$. Figure 4.10 shows this situation. This yields $\Delta_{2b} = \frac{1}{4(k+1)}(k+1) = \frac{1}{4}$, if a tower of length k starts in f in $\vec{Y}$ and $\Delta_{2b} = 0$ otherwise. So for this case in total the change in distance can be bound by

$$\Delta_{2b} \le \frac{1}{4}.$$

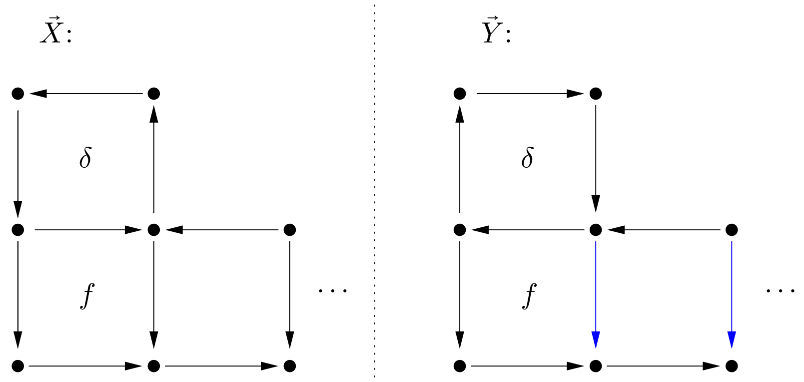

Figure 4.10: Situation 2.b, f is scrambled in $\vec{X}$ and might start a tower in $\vec{Y}$.

c) If f is blocked in one of the two instances, it is either scrambled or oriented in the other one. Therefore, this case is already covered by the previous ones.

Combining the three cases, we have

$$\Delta_2 \leq \tfrac{1}{4}.$$

3. For $f \neq \delta$ and f and δ do not share an edge, we again have some different cases to consider:
   a) If f is oriented in $\vec{X}$ or $\vec{Y}$, it is oriented in both in the same direction, so the chain applies the same operation to both orientations and the distance does not change at all, i.e.
   $$\Delta_{3a} = 0.$$
   b) If f starts a tower in $\vec{X}$ or $\vec{Y}$, which has no edge in common with δ, the same tower exists in the other orientation and the step of the coupling does not change the distance, because both instances behave in the same way. This leads to
   $$\Delta_{3b} = 0.$$
   c) If a tower $T = (a_k, \ldots, a_0)$ starts in f (without loss of generality in $\vec{X}$ and oriented cw), which means $a_k = f$, which does touch δ. So according to Lemma 11, the edge of a_0 opposing the blocking edge of a_1 is the edge of δ, which they have in common. This implies $(f, a_{k-1}, \ldots, a_1, \delta)$ is a tower of length $k+1$ in $\vec{Y}$. Also note, that both towers are oriented in the same direction, as Figure 4.11 shows. This yields, for $p \leq 1 - \frac{1}{4(k+1)}$ no orientation is changed by the coupling step. For $1 - \frac{1}{4(k+1)} < p \leq 1 - \frac{1}{4(k+2)}$ only the tower in $\vec{X}$ is flipped, which increases the distance by $k+1$. Finally, for $p \geq 1 - \frac{1}{4(k+1)}$ both towers are flipped, which decreases the distance by 1. In total:
   $$\Delta_{3c} = \left(\frac{1}{4(k+1)} - \frac{1}{4(k+2)}\right) \cdot (k+1) + \frac{1}{4(k+2)} \cdot (-1) = 0$$

   So in total, case 3 leads to
   $$\Delta_3 = 0.$$

Now, it remains to merge all cases, weighted accordingly to the number of faces, which can result in the corresponding case. There is one case of type 1, four cases of type 2 and a $n-7-5$ faces of cases of type 3. So the total expected distance after one step of the coupling P on a set of orientations at distance 1 is

$$\Delta \leq 1 \cdot (-1) + 4 \cdot \tfrac{1}{4} + (n-7-5) \cdot 0 = 0$$

resulting in

$$\mathbb{E}_P[d(X_{t+1}, Y_{t+1}) \mid (\vec{X}, \vec{Y}) = (X_t, Y_t)] \leq 1,$$

which means $\beta = 1$ complies with all the conditions, required by Theorem 15.

□

So after this prelude, finally the application of Theorem 15 is possible, since all the building blocks, which are needed, were finally established.

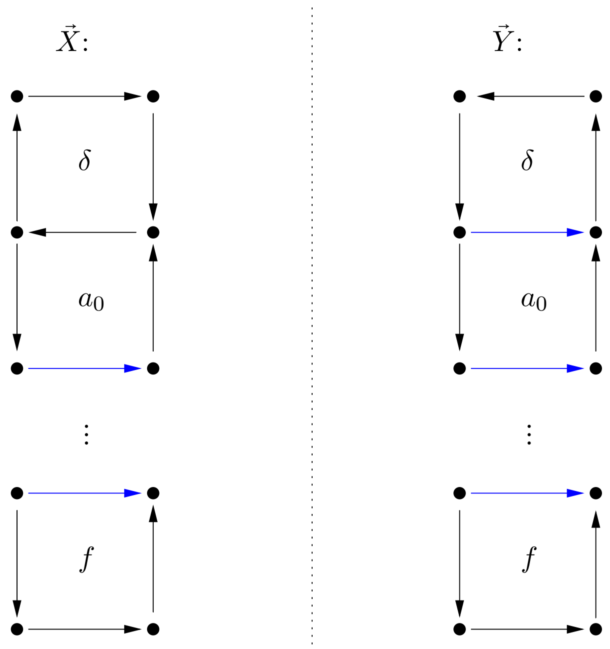

Figure 4.11: Situation 3c, a tower in $\vec{X}$ touching δ implies a tower in $\vec{Y}$ ending in δ.

Proposition 4 *(mixing time of $M_T^\star$)*
Let $\mathcal{M}_T$ be the transition matrix of the tower Markov chain M_T and $M_T^\star$ the Markov chain with transition matrix $\mathcal{M}_T^\star := \frac{1}{1+\widetilde{p}}(\mathcal{M}_T + \widetilde{p}\cdot\mathcal{I})$, where $\mathcal{I}$ denotes the $\Omega\times\Omega$ identity matrix and $\widetilde{p} = \frac{1}{2(n-7)}$ as established in Lemma 13. The mixing time $\tau_{M_T^\star}$ of the Markov chain $M_T^\star$ is bounded by

$$\tau_{M_T^\star}(\varepsilon) \le \frac{e}{2}\cdot n^5 \cdot \log(\varepsilon^{-1}).$$

Proof. As the previous three Lemmas have shown, Theorem 15 is applicable with $\beta = 1$, $p = \frac{1}{2(n-7)}$, and $D \le \frac{n^2-4n+4}{2} < \frac{n^2}{2}$ to our coupling. This yields:

$$\begin{aligned}
\tau_{M_T^\star}(\varepsilon) &\le \left\lceil \frac{eD^2(1+\widetilde{p})}{\widetilde{p}} \right\rceil \cdot \left\lceil \log(\varepsilon^{-1}) \right\rceil \\
&\le \frac{e(\frac{n^4}{4})(1+\frac{1}{2(n-7)})}{\frac{1}{2(n-7)}} \cdot \log(\varepsilon^{-1}) \\
&\le \frac{e}{2}\cdot n^5 \log(\varepsilon^{-1})
\end{aligned}$$

□

Now it remains to translate this bound to the mixing time of $M_T^\star$ to a bound of the mixing time of M.

Proof. (of Theorem 19)
We want to apply Theorem 16 to bound the mixing time τ_M of M in terms of $\tau_{M_T^\star}$, which was bounded in Proposition 4. All Markov chains M, M_T, and $M_T^\star$ are reversible and ergodic, so they share as stationary distribution π the uniform distribution. Further,

all three chains are lazy, i.e. with a probability of $\geq \frac{1}{2}$ they do not change an element, so their second largest eigenvalue is $\geq \frac{1}{2}$. This means Theorem 16, the comparison theorem of Randall and Tetali, can be applied. To do so, we have to translate all moves of $M_T^\star$ in paths of moves of M and bound

$$A := \max_{(x,y)\in E(M)} \left\{ \frac{1}{\pi(x)\cdot(\mathcal{M})_{(x,y)}} \cdot \sum_{(u,v)\in\Gamma(x,y)} |\gamma_{uv}| \cdot \pi(u) \cdot (\widetilde{\mathcal{M}})_{x,y} \right\}$$

where

$$\Gamma(x,y) := \{(u,v) \in E(\widetilde{M}) \mid (x,y) \in \gamma_{uv}\}.$$

The steps we have to consider are on the one hand, flips of towers in $M_T^\star$, which can be rephrased as a sequence of flips of directed faces, see Corollary 6. So let $(u,v) \in E(M_T)$ be a flip of a tower of length k. There is a natural sequence of face flips $\gamma_{\vec{U}\vec{V}}$ of length $k+1$ of flips of faces, which results in the same total change. This leads to

$$\begin{aligned} |\gamma_{\vec{U}\vec{V}}| \cdot \pi(\vec{U}) \cdot (\mathcal{M}_T)_{\vec{U},\vec{V}} &= (k+1)\cdot \frac{1}{|\Omega|} \cdot \frac{1}{4(k+1)} \cdot \frac{1}{(n-7)} \\ &= \frac{1}{|\Omega|} \cdot \frac{1}{4(n-7)}, \end{aligned}$$

where $(\mathcal{M}_T)_{\vec{U},\vec{V}}$ is the entry of the transition matrix $\mathcal{M}_T$ of the Markov chain M_T, associated with the orientations $\vec{U}$ and $\vec{V}$. In other words, it is the transition probability between the two orientations, $(\mathcal{M}_T)_{\vec{U},\vec{V}} = \Pr(X_{t+1} = \vec{V} \mid X_t = \vec{U})$. Further, the equations hold, because there is one of $n-7$ flippable faces, which we have to choose from and we flip the oriented tower with probability $\frac{1}{4(k+1)}$. Also, there are at most $4(n-7)$ towers which contain a face f, because at any of the $n-7$ flippable faces could start four towers containing the face f. So at most $4(n-7)$ ways γ_{uv} contain a fixed face flip.
Note, that the family of examples in Remark 5 shows, that this bound is tight up to a constant factor of 4, since in any face of one of the graphs, a tower could start, which ends in the middle face.
Further, we know for a fixed pair $(\vec{X},\vec{Y}) \in E(M) = S$ with $\vec{X} \neq \vec{Y}$

$$(\mathcal{M}_T^\star)_{\vec{U},\vec{V}} = \tfrac{1}{1+p}(\mathcal{M}_T + p\mathcal{I})_{\vec{U},\vec{V}} \leq (\mathcal{M}_T)_{\vec{U},\vec{V}}.$$

This yields

$$\begin{aligned} A_{\vec{X},\vec{Y}} &:= \frac{1}{\pi(\vec{X})\cdot(\mathcal{M})_{\vec{X},\vec{Y}}} \cdot \sum_{(\vec{U},\vec{V})\in\Gamma(\vec{X},\vec{Y})} |\gamma_{\vec{U}\vec{V}}| \cdot \pi(\vec{U}) \cdot (\mathcal{M}_T^\star)_{\vec{U},\vec{V}} \\ &\leq \frac{1}{\pi(\vec{X})\cdot(\mathcal{M})_{\vec{X},\vec{Y}}} \cdot \sum_{(\vec{U},\vec{V})\in\Gamma(\vec{X},\vec{Y})} |\gamma_{\vec{U},\vec{V}}| \cdot \pi(\vec{U}) \cdot (\mathcal{M}_T)_{\vec{U},\vec{V}} \\ &\leq \frac{1}{\frac{1}{|\Omega|}\cdot\frac{1}{2(n-7)}} \cdot 4(n-7) \cdot \frac{1}{|\Omega|} \cdot \frac{1}{8(n-7)} \\ &= (n-7), \end{aligned}$$

which does not depend on $\vec{X}$ or $\vec{Y}$.

This is not applicable to pairs $(\vec{X},\vec{X})$, so we have to take them into account now: First, $\Gamma(\vec{X},\vec{X}) = \{(\vec{X},\vec{X})\}$, since the only path in the transitions of $M_T^\star$ using the loop, is the same loop in M. So

$$\begin{aligned}
A_{\vec{X}\vec{X}} &= \frac{1}{\pi(\vec{X})\cdot(\mathcal{M})_{\vec{X},\vec{X}}}\cdot\sum_{(\vec{U},\vec{V})\in\Gamma(\vec{X},\vec{Y})}|\gamma_{\vec{U},\vec{V}}|\cdot\pi(\vec{U})\cdot(\mathcal{M}_T^\star)_{\vec{U},\vec{V}}\\
&= \frac{1}{\pi(\vec{X})\cdot(\mathcal{M})_{\vec{X},\vec{X}}}\cdot\pi(\vec{X})\cdot(\mathcal{M}_T^\star)_{\vec{X},\vec{X}}\\
&= \frac{1}{\pi(\vec{X})\cdot(\mathcal{M})_{\vec{X},\vec{X}}}\cdot\pi(\vec{X})\cdot\tfrac{1}{1+p}(\mathcal{M}_T+p\mathcal{I})_{\vec{X},\vec{X}}\\
&\le \frac{1}{\pi(\vec{X})\cdot(\mathcal{M})_{\vec{X},\vec{X}}}\cdot\pi(\vec{X})\cdot\tfrac{1}{1+p}\\
&\le \frac{1}{(\mathcal{M})_{\vec{X},\vec{X}}}\\
&\le 2,
\end{aligned}$$

since $\frac{1}{2}\le(\mathcal{M})_{\vec{X},\vec{X}}\le 1$. Together, this directly yields $A = \max_{(\vec{X},\vec{Y})\in E(M)} A_{\vec{X}\vec{Y}} \le n-7$.

Finally, we need a simple upper bound to $\pi^\star = \min_{\vec{X}\in\Omega}\pi(\vec{X}) = \frac{1}{|\Omega|}$. Note, the graph has n vertices, so it contains $2(n-4)$ edges. This leads to the trivial upper bound $|\Omega|\le 2^{2n-4} < 2^{2n}$, if all orientations of the graph, not only 2-orientations, are taken into account. This implies $\frac{1}{\pi_\star} = |\Omega| \le 2^{2n}$. Last but not least we have all parts together and Theorem 16 yields:

$$\begin{aligned}
\tau_M(\varepsilon) &\le \frac{4\log(\frac{1}{\varepsilon\cdot\pi_\star})}{\log(\frac{1}{2\varepsilon})}\cdot A\cdot\tau_{M_T^\star}(\varepsilon)\\
&= \frac{4\log(\frac{2^{2n}}{\varepsilon})}{\log(\frac{1}{2\varepsilon})}\cdot 2(n-7)\cdot\tau_{M_T^\star}(\varepsilon)\\
&\le \frac{4(2n+\log(\varepsilon^{-1}))}{\log(\frac{1}{2\varepsilon})}\cdot 2(n-7)\cdot\frac{e}{2}\cdot n^5\cdot\log(\varepsilon^{-1})\\
&\le \frac{4e(2n+\log(\varepsilon^{-1}))}{\log(\frac{1}{2\varepsilon})}\cdot n^6\cdot\log(\varepsilon^{-1}).
\end{aligned}$$

□

4.4 α-Orientations of plane triangulations

This last section treats some one aspect of the behaviour of the face flip Markov chain M for general α-orientations of plane triangulations. As established earlier, for graphs with a high maximum degree, the face flip Markov chain M cannot be expected to be rapid mixing. But even worse, there is a family of plane triangulations of maximum degree 6, where the chain torpidly mixes the set of α-orientations for every of these graphs. This family is built from the graphs G_n, which were used in Theorem 17. Taking the graph G_{3k-1}, the vertex x_0 is replaced by a token graph T_k, which is defined as follows:

Definition 24 (the token graph and the weight function $\widetilde{\alpha}$)
The token T_0 consists of one single vertex c_0.
Next, T_1 consists of T_0 and the six-cycle $c_1, b_{1,1}, b_{1,2}, c_{-1}, d_{1,2}, d_{1,1}$. Further, all vertices of the six-cycle are adjacent to c_0. So we have a hexagonal subset of the triangular lattice of side-length 1 with a middle vertex c_0.

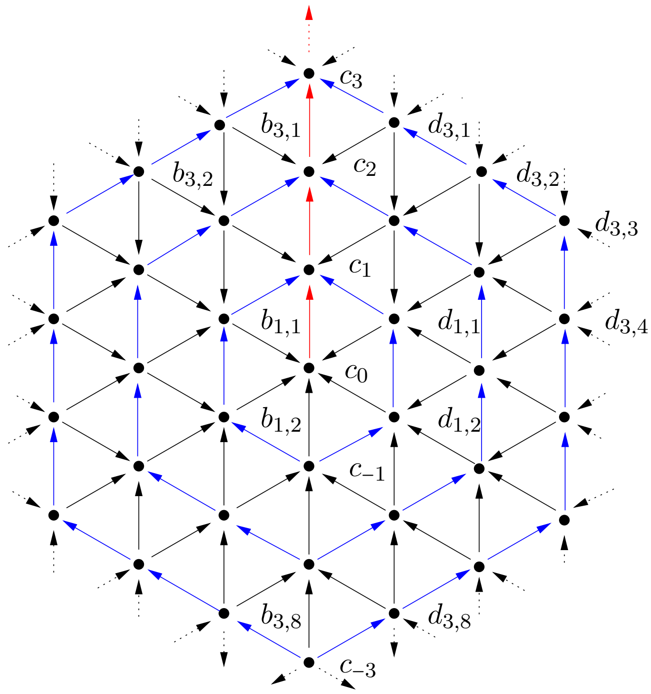

Figure 4.12: The token graph T_3, with red outgoing path from c_0, blue flippable cycles and the dashed in- and outgoing demand.

In general, the token T_k consists of T_{k-1} and $6k$ additional vertices $c_k, b_{k,1}$, $b_{k,2}$, …, $b_{k,3k-1}$, c_{-k}, $d_{k,3k-1}$, …, $d_{k,1}$, which form one cycle (in the order above) and are connected to the vertices of T_{k-1} such that c_{k-1} is connected to $b_{k,1}$, c_k, and $d_{k,1}$ and $d_{k-1,1}$ is connected to $d_{k,1}$ and $d_{k,2}$ and T_k is a hexagonal subset of the triangular grid with corners $c_k, b_{k,k}, b_{k,2k}, c_{-k}, d_{k,2k}, d_{k,k}$ with side length k. This implies there are in total 12k-6 edges, connecting T_{k-1} to the new cycle, but they are fixed by the five edges above and the triangular grids structure.
Besides the token graph, there is need for the definition of the weight-function $\widetilde{\alpha}$, which specifies the out-degrees of the vertices of T_k. This function $\widetilde{\alpha}$ is, as follows:

- $\widetilde{\alpha}(c_i) := 1$ for $i = 0, 1, \dots, k$
- $\widetilde{\alpha}(c_i) := 3$ for $i = -1, -2, \dots, -(k-1)$ and $\widetilde{\alpha}(c_{-k}) := 5$

- $\widetilde{\alpha}(b_{i,j}) := 3$ for $j \notin \{i, 2i\}$ and $(i,j) \neq (k, 3k-1)$
- $\widetilde{\alpha}(b_{i,j}) := 2$ if $j = i$ or $j = 2i$ for $i = 1, \ldots, k$
- $\widetilde{\alpha}(d_{i,j}) := 3$ for $j \notin \{i, 2i\}$ and $(i,j) \neq (k, 3k-1)$
- $\widetilde{\alpha}(d_{i,j}) := 2$ if $j = i$ or $j = 2i$ for $i = 1, \ldots, k$
- $\widetilde{\alpha}(b_{k,3k-1}) := 4$ and $\widetilde{\alpha}(d_{k,3k-1}) := 4$

Note, $c_k, b_{k,k}, b_{k,2k}, c_{-k}, d_{k,2k}, d_{k,k}$ are the corners of the hexagon, which have to be treated slightly different regarding $\widetilde{\alpha}$. One such orientation is depicted in Figure 4.12 for T_3, but it can easily be extended to all k.

To understand, why this token has the right properties, consider the role of the vertex x_0 in the graphs G_n of Theorem 17, which are drawn in Figure 4.2. There $\deg(x_0) = 2n+3$. Further, it has 2 rigid outgoing edges to a and b, which are vertices on the outer face, which leaves a non-rigid outgoing edge. Finally, any of the non rigid $2n+1$ edges of x_0 can be the third outgoing one. In this section, x_0 is going to be replaced by T_k, so T_k should behave in the same way as x_0 does. Before this is proven, we have to exemplify, now T_k replaces x_0:

Definition 25 (the graph $\widetilde{G}_k$)
Let $\widetilde{G}_k$ be the graph G_{3k-1}, where the vertex x_0 is substituted by the token graph T_k in such a way, that c_k is connected to v_1, x_1 and w_1. Further, $b_{k,i}$ is connected to v_i and v_{i+1} as well as $d_{k,i}$ to w_i and w_{i+1} for all $i = 1, \ldots, 3k-1$. Finally, c_{-k} is connected to the vertices a and b of the outer face. Also $\{b_{k,3k-1}, a\}$ and $\{d_{k,3k-1}, b\}$ are edges of $\widetilde{G}_k$.

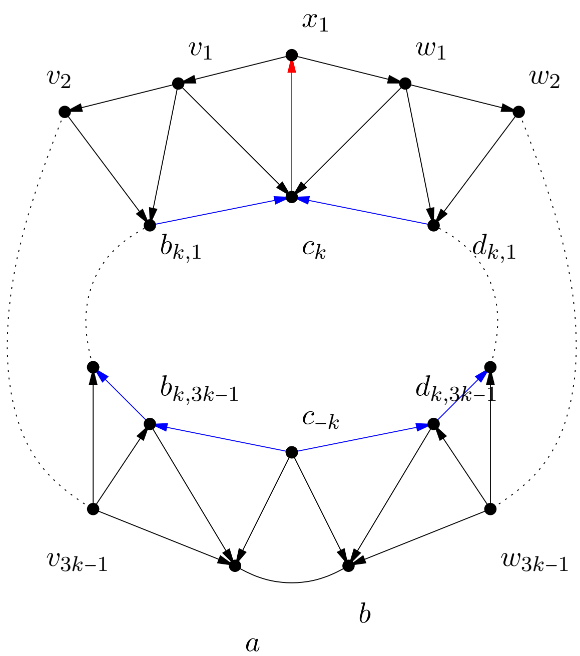

Figure 4.13: Connecting T_k and G_{3k-1}.

This substitution is presented in Figure 4.12, while Figure 4.14 shows the graph $\widetilde{G}_2$. Every vertex v_i and w_i of G_{3k+1} is connected to two vertices of T_k, so they all have a degree of 6. Further, x_1 has one more edge and so $\deg(x_1) = 6$. Also all vertices $b_{k,i}$ and $d_{k,i}$ have two additional edges, so they have a degree of 5 (if $i = k$ or $i = 2k$) or 6. The vertex c_k is adjacent to three vertices in G_{3k-1}, so it also has a degree of 6, while c_{-k} ends up with a degree of 5. Since all other vertices do not receive any additional edges, they have a degree of at most 6 and $\widetilde{G}_k$ has therefore a maximum degree of 6.

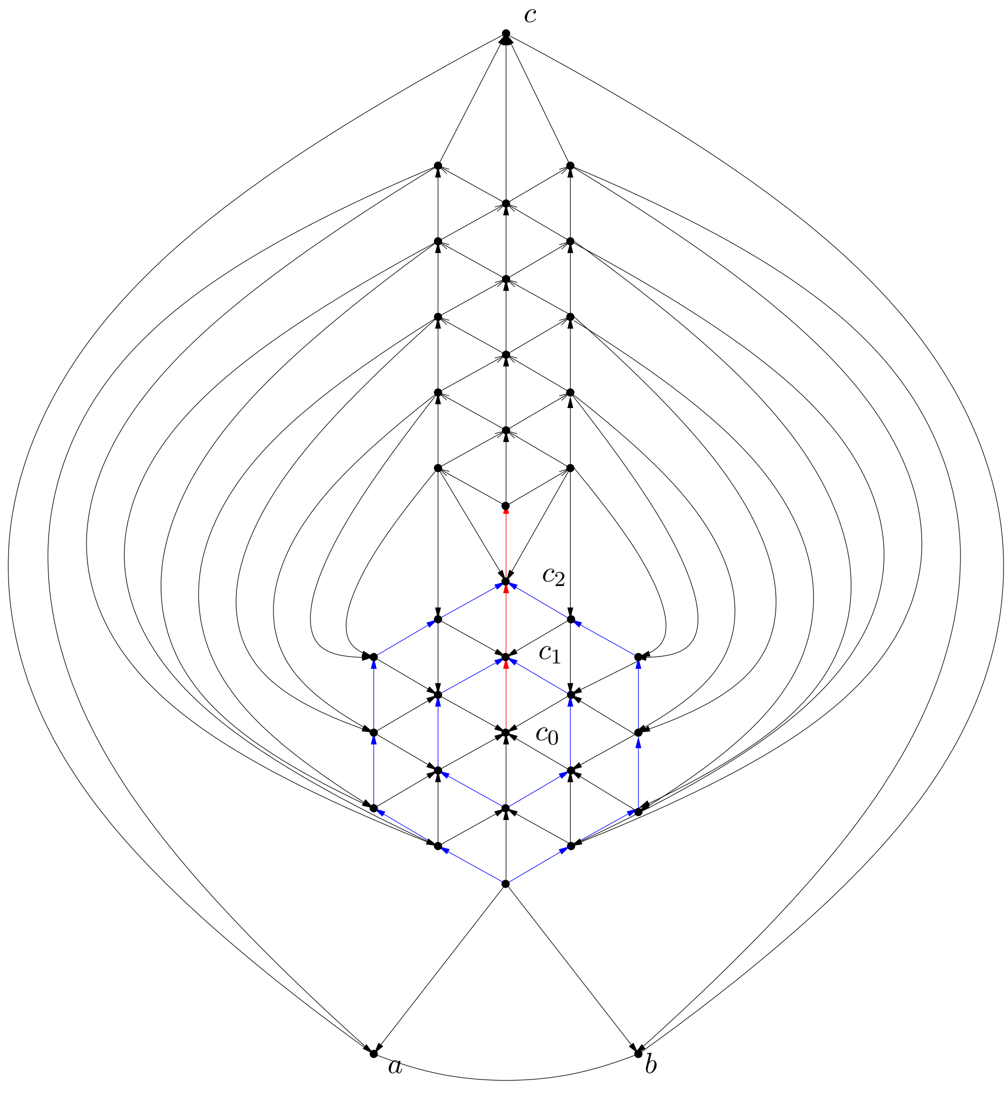

Figure 4.14: The 'green' orientations of $\widetilde{G}_2$, which consists of T_2 properly inserted into G_5.

Since we already defined $\widetilde{\alpha}$ for all vertices of T_k, we now extend it to $\widetilde{G}_k$. We prescribe out degrees of the vertices of G_{3k-1} as follows

- $\widetilde{\alpha}(x_i) = 3$ for $i = 1, \ldots, 3k-1$,
- $\widetilde{\alpha}(v_i) = 4$ for $i = 1, \ldots, 3k-2$ and $\widetilde{\alpha}(v_{3k-1}) = 5$, and

- $\widetilde{\alpha}(w_i) = 4$ for $i = 1, \dots, 3k-2$ and $\widetilde{\alpha}(w_{3k-1}) = 5$
- $\widetilde{\alpha}(a) = 0$, $\widetilde{\alpha}(b) = 1$, and $\widetilde{\alpha}(c) = 2$

and get a weight function $\widetilde{\alpha}$ for all vertices of $\widetilde{G}_{3k-1}$.

As announced, if we insert T_k as we do it in $\widetilde{G}_k$, it has –more or less– the same properties as x_0 has in G_{3k-2}:

Lemma 15 *(Properties of a properly inserted T_k)*
Let T_k be inserted into some graph as shown in Figure 4.13. The token T_k has the following properties:

1. *T_k is connected with $2(3k-2)+3$ edges to the outside, five of them outgoing*
2. *T_k is connected with 4 rigid outgoing edges to the outer face, two edges to a and two to b*
3. *there is only one $\widetilde{\alpha}$-orientation of T_k, such that this path leaves via the middle edge of c_k*

Proof.

1. The outer circle of T_k consists of $6k$ vertices and $2 \cdot 6k + 1$ edges to the outside. One vertex 'on the other side' is adjacent to only one of these edges, every other one is adjacent to exactly to edges. So the size of the neighbourhood of T_k, the number of vertices, T_k is connected to, is $6k - 1 = 2(3k-2) + 3$.

   Further, the vertices outer circle of T_K, namely $c_k, c_{-k}, d_{k,j}$ and $b_{k,j}$ need in total $1+5+6(k-1)3+2+4\cdot 2 = 18k-2$ outgoing edges. They are connected in between with $6k$ edges. Further, there are $(6(k-1)2+6$ edges from the outer circle of T_k to its inner T_{k-1}, of whom $(12(k-1)+5$ are outgoing of the outer circle. So there are $18k-7$ edges outgoing of the outer circle, leaving a demand of 5.

2. T_k is connected to a and b on the outer face. Each of these vertices is connected by 2 edges, so T_k has 4 rigid outgoing edges to the outer face.

3. If the edge $\{c_k, x_1\}$ is oriented towards x_1, the vertex c_k is saturated, because $\widetilde{\alpha}(c_k) = 1$. This forces $\{c_{k-1}, c_k\}$ to be oriented towards c_k, which saturates c_{k-1}. Inductively all edges $\{c_{i-1}, c_i\}$ for $i = 1, \dots, k$ are oriented towards c_i and all c_i for $i \geq 0$ are saturated.

   So if c_i is saturated, the edges $\{d_{i,1}, c_i\}$ and $\{b_{i,1}, c_i\}$ have to point towards c_i. Since the one free outgoing edge of T_{i-1} goes to c_i, the edges connecting $d_{i,1}$ and $b_{i,1}$ with T_{i-1} are oriented towards T_{i-1}, which saturates $d_{i,1}$ and $b_{i,1}$. Now, inductively, all other edges $\{d_{i,j-1}, d_{i,j}\}$ and $\{b_{i,j-1}, b_{i,j}\}$ have to be oriented towards $d_{i,j-1}$ (or $b_{i,j-1}$ respectively). This leaves no choice, than to orient every edge as shown in Figure 4.12.

   Another way to argue is again based on Figure 4.12. As one can see, there is no oriented cycle to flip in the shown orientation, which does not contain (c_k, x_1), so any other $\widetilde{\alpha}$-orientation of T_k contains this edge in the opposing direction.

□

This lemma could even be extended, since for any orientation of a properly inserted T_k, we have exactly one directed path starting in c_0 leaving the token somewhere. There might be multiple orientations, such that the path leaves the token at the same vertex, but there is always exactly one such path, which can easily be constructed, starting with the free edge which leaves T_k and than tracing back.
Now, having established the family of graphs to consider and our the prescribed out-degrees $\widetilde{\alpha}$, it is time to show, that the set of $\widetilde{\alpha}$-orientations of $\widetilde{G}_k$ is hour glass-shaped, which implies, that the face flip Markov chain M mixes $\Omega_{\widetilde{G}_k,\widetilde{\alpha}}$ torpidly:

Theorem 20
The face flip Markov chain does not rapidly mix the set of $\widetilde{\alpha}$-orientations of any graph of the infinity family of planar triangulations $\{\widetilde{G}_k \mid k \in \mathbb{N}, k > 0\}$, i.e. the mixing time of the face flip Markov chain M on the set of $\widetilde{\alpha}$-orientations of $\widetilde{G}_k$ is bounded by

$$\tau_M \geq 2^{3k-4} - 1.$$

Proof. The line of arguments, which is going to be applied here is rather obvious, since it was already made use of twice in this work, namely in Theorem 17 and Theorem 18. So let k be fixed, consider the graph $\widetilde{G}_k$, and let Ω be the set of its $\widetilde{\alpha}$-orientations. To show, that Ω is hour-glass-shaped, let

$$\begin{aligned} \Omega_g &:= \{\vec{G} \in \Omega \mid (c_k, x_1) \in E(\vec{G}\}, \\ \Omega_r &:= \{\vec{G} \in \Omega \mid \exists i,j : (b_{k,j}, v_i) \in E(\vec{G})\}, \text{ and} \\ \Omega_b &:= \{\vec{G} \in \Omega \mid \exists i,j : (d_{k,j}, w_j) \in E(\vec{G})\}. \end{aligned}$$

So in Ω_r the outgoing edge of T_k is left of (x_1, c_k) and Ω_b contains all orientations, where it is on the right hand side. As previously, the following arguments make sure, the following three claims are met:

- Claim $|\Omega_g| = 1$: In a 'green' orientation the edge $\{c_k, x_1\}$ is directed from T_k to G_{3k-1}. So according to Lemma 15, (3), there is only one such orientation of T_k. Further, there is only one 'green' orientation of G_{3k-1}, if it contains (x_0, x_1), which combined yields this claim.
- Claim $|\Omega_r| = |\Omega_b| \geq 2^{3k-2}$: As in the previous proofs, the symmetry of $\widetilde{G}_k$ yields this claim. Flipping an orientation of $\widetilde{G}$ at the c_{-k}, $c_{-(k-1)}$, $\dots$, c_0, $\dots$, c_k, x_1, $\dots$, x_{3k-1} - axis, we find a bijection between Ω_r and Ω_b, as exposed twice earlier.

  Consider the green orientation of $\widetilde{G}_k$. There the cycle c_k, x_1, v_1, $\dots$, v_{3k-1}, $d_{k,3(k-1)}$, $c_{-(k-1)}$, $c_{-(k-2)}$, $\dots$, c_{k-1} in the green orientation is oriented and can be reversed. This yields a red orientation, where the triangles v_i, v_{i+1}, x_i for $i = 1, \dots, (3k-1) - 1$ are all oriented and can be flipped independently of each other. So there are at least 2^{3k-2} red orientations.
- Claim: Ω_g separates Ω_b from Ω_r: In any $\widetilde{\alpha}$-orientation of $\widetilde{G}_k$ there is exactly one edge pointing from T_k to the vertices of G_{3k-1}, except the rigid edges to a and b, which are on the outer face. Assume $\vec{G} \in \Omega_r$, so this edge goes to some v_i. For any blue orientation, it has to point to some w_j, so we have to flip it, step by

step from one side to the other. This can only be done using the edge $\{c_k, x_1\}$, because on the other side of T_k we have rigid edges from c_{-k} to a and b, which prevent going this path. So on any path from an element of Ω_r to an element of Ω_b the outgoing green edge of T_k has to be (c_k, x_1), which means, that this path hits Ω_g.

So Ω is an hour glass-shaped state lattice. Therefore the requirements of Lemma 9 are met and it can be applied:

$$\Phi_M \leq \frac{\pi(\Omega_g)}{\pi(\Omega_b)} = \frac{|\Omega_g|}{|\Omega|} \cdot \frac{|\Omega|}{|\Omega_b|} \leq \frac{1}{2^{3k-2}}$$

and Theorem 14 leads to

$$\begin{aligned} \tau_M &\geq \frac{1}{4 \cdot \Phi_M} - \frac{1}{2} \\ &\geq \frac{2^{3k-2}}{4} - \frac{1}{2} \\ &\geq 2^{3k-4} - 1, \end{aligned}$$

which completes this proof.

□

This section ends with a remark, which might improve the graph $\widetilde{G}_k$ slightly, depending on the readers opinion, what a nice graph is:

Remark 6
The vertex c_{-k} of $\widetilde{G}_k$ is connected to 5 rigid edges. It is possible to delete c_{-k} from $\widetilde{G}_k$ and all of its adjacent edges and insert edges from $b_{k-1,3(k-1)-1}$ and $d_{k-1,3(k-1)-1}$ to a. Since a rigid sub graph is replaced by another rigid one, the resulting graph has the same set of $\widetilde{\alpha}$-orientations (except we have to increase $\widetilde{\alpha}(b_{k-1,3(k-1)-1})$ and $\widetilde{\alpha}(d_{k-1,3(k-1)-1})$ by one, since they both got an additional rigid, outgoing edge to a). The resulting graph looses its symmetry, but does not have any rigid edges, except the ones which go towards vertices on the outer face.

4.5 Conclusion

We would like to close this capter with some open questions:

- Is there an analogue of Remark 4 for 2-orientations of plane quadrangulations of maximum degree 4? Even if not, is it really necessary to use the theorem of Bordewich and Dyer instead of the theorem of Dyer and Greenhill? Remark 4 does not imply, that the Dyer Greenhill is not applicable, but you have to argue in another manner and you will have to include tower-moves in the argument, to be able to apply the theorem. To us it is currently unclear, if this is possible or not.

- Is there an analogue of Corollary 5 to Theorem 17? This means, are we able to find a family of plane quadrangulations with a maximum degree in the order of $\sqrt{\#\text{vertices}}$, such that the face flip Markov chain is not rapidly mixing?

- Is it possible to improve the maximum degree of Theorem 19 or its analogue for plane quadrangulations of maximum degree 6 in [39]? In both cases there is still quite a bit gap between 6 and $\mathcal{O}(\sqrt{\#\text{vertices}})$ and respectively 4 and $\mathcal{O}(\#\text{vertices})$. We think, it is rather unclear, what the order of the mixing time of the face flip Markov chain in between is.

5 Conclusion

We already saw some conclusions, which were part of Chapter 2, 3, and 4. Now, let us look briefly at the complete picture:

- In Chapter 2 we introduced the method of block coupling and applied it to various families of graphs, forming sets of 2-heights. Especially, we sat that the up/down-Markov chain is rapidly mixing the sets of 2-heights of
  - toroidal triangle grids,
  - triangle grid graphs,
  - toroidal rectangular grid graphs, and
  - plane triangulations.

  Still, we failed providing results for $k = 3$ or even bigger k, due to the massive amount of computations for slightly bigger blocks. The original intent of Chapter 3 was to understand the behaviour of k-heights on a path well enough and apply this knowledge to decrease the required amount of computations. We were able to find some formulas which explain well the behaviour of these k-heights, or bounded lattice paths, but did not make the way back to the general block coupling of k-heights. So the first open question is: *Can the explicit formulas for bounded lattice paths be applied to speed up the computation of $E_{B,\delta}$ to produce results for 3-heights or even bigger k?* Also, there are some interesting open questions next to the block coupling method, which we already stated:
  - First, obviously there is the question if the coefficients, which occur in the bound of the mixing time established with the block coupling methods are required to be that huge. It might be worth it to improve them. Some first fooling around with coupling from the past indicated much better mixing times. *These experiments could be substantiated and the method itself might be improveable in terms of the resulting bounds to the mixing time.*
  - Second, the block coupling method does not guide you, what good blocks are. Some more insight into what makes a set of blocks good would be very delightful. Currently, it is not even clear if it is desirable to create blocks with small or with big border (compared to the amount of interior faces). We tried some different shapes of blocks, but did not gain experimental insight.
- With Chapter 3 we did some detour from Markov chains into lattice path enumeration. We saw explicit formulas for some families of bounded lattice paths, which are somewhat the *one dimensional* instance of the two dimensional problem of investigating the up/down Markov chain on planar graphs. Maybe these insights might help to understand, what good blocks look like. Maybe it might even be possible to use this to improve the required amount of computations to decide if

a set of blocks is applicable to the block coupling method. We failed do to so, maybe somebody else proceeds here. Also, the method used to count the bounded lattice paths are worthwhile itself, so continued study and improvement of them should be considered.

- The overall question in the first place was to investigate the behaviour of the face flip Markov chain on α-orientations with the block coupling method. This is missing in this work. Still some progress on the behaviour of this chain was presented in Chapter 4 using tower moves as coupling method for positive results and conductance for negative ones. Especially, we were able to present
  - a family of planar triangulations of maximum degree $\sqrt{n}$ with $O(n)$ vertices, such that the face flip Markov chain is not rapidly mixing the set of 3-orientations of these graphs.
  - a family of planar quadrangulation of maximum degree n with $O(n)$ vertices such that the face flip Markov chain is not rapidly mixing the set of 2-orientations of these graphs.
  - a proof, stating that for planar quadrangulations of maximum degree 4 the chain is indeed rapidly mixing.
  - Presenting a family of triangulations of maximum degree 7, such that for an α (with $\alpha(v) = 3$ for most vertices v) the face flip Markov chain is not rapid lymixing the set of α-orientations.

  In the end, the gap of families of graphs for which the chain might be rapidly mixing has been reduced to quite some extend by this work. *Still it might be worthwhile to repeat the tower moves results with block coupling and try to extend them.*

So, as usual at the end of some mathematical work, some work was done and even more revealed to be considered in future...

List of Figures

Bibliography

[1] *7th international conference on lattice path combinatorics and applications*, July 2010, (`http://www.unisi.it/eventi/lattice_path_2010/`).

[2] Rudolf Ahlswede and David E. Daykin, *An inequality for the weights of two families of sets, their unions and intersections*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **43** (1978), no. 3, 183–185.

[3] Benjamin Armbruster, *personal communication*, 2013.

[4] ———, *A short proof of Strassen's theorem*, 2013, (`http://www.optimization-online.org/DB_FILE/2013/08/4002.pdf`).

[5] Cyril Banderier and Pierre Nicodeme, *Bounded Discrete Walks*, Proceedings of the AofA 2010 Conference (Vienna, Autria) (DMTCS, ed.), June 2010, pp. 35–48.

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

[7] J. Berg and R. Brouwer, *Random sampling for the monomer-dimer model on a lattice*, Tech. report, Amsterdam, The Netherlands, 1999.

[8] Magnus Bordewich and Martin Dyer, *Path coupling without contraction*, Journal of Discrete Algorithms **5** (2007), 280–292.

[9] Albrecht Böttcher and Sergei M. Grudsky, *Spectral properties of banded Toeplitz matrices.*, SIAM, 2005.

[10] Timothy Chow and Julian West, *Forbidden subsequences and chebyshev polynomials.*, Discrete Mathematics **204** (1999), no. 1-3, 119–128.

[11] G.M. Cicuta, M. Contedini, and L. Molinari, *Enumeration of simple random walks and tridiagonal matrices.*, J. Phys. A, Math. Gen. **35** (2002), no. 5, 1125–1146.

[12] Páidí Creed, *Sampling Eulerian orientations of the triangular lattice graphs*, Disc. Alg. **7** (2009), no. 2, 168–180.

[13] Hubert de Fraysseix and Patrice Ossona de Mendez, *On topological aspects of orientations.*, Discrete Mathematics **229** (2001), no. 1-3, 57–72.

[14] Hubert de Fraysseix, Patrice Ossona de Mendez, and Pierre Rosenstiehl, *Bipolar orientations revisited.*, Discrete Applied Mathematics **56** (1995), no. 2-3, 157 – 179.

[15] Persi Diaconis and Laurent Saloff-Coste, *Comparison theorems for reversible Markov chains*, Annals of Applied Probability **3** (1993), 696–730.

[16] Murray Dow, *Explicit inverses of Toeplitz and associated matrices*, ANZIAM J. **44** (2003), no. E, E185–E215.

[17] Martin Dyer and Catharine Greenhill, *A more rapidly mixing Markov chain for graph colourings*, Random Structures and Algorithms **13** (1998), 285–317.

[18] Johannes Fehrenbach and Ludger Rüschendorf, *Markov chain algorithms for Eulerian orientations and 3-colourings of 2-dimensional cartesian grids*, Statistics & Decisions **22** (2004), 109–130.

[19] William Feller, *An introduction to probability theory and its applications*, vol. 1, Wiley, January 1968.

[20] Stefan Felsner, *Lattice structures from planar graphs*, Electron. J. Comb. **11** (2004), no. 1, 24p.

[21] ______, *Rectangle and square representations of planar graphs.*, Pach, János (ed.), Thirty essays on geometric graph theory. Berlin: Springer. 213-248 (2013)., 2013.

[22] Stefan Felsner, Éric Fusy, Marc Noy, and David Orden, *Bijections for Baxter families and related objects*, J. Comb. Theory Ser. A **118** (2011), no. 3, 993–1020.

[23] Stefan Felsner and Daniel Heldt, *Lattice path enumeration and Toeplitz matrices*, Journal of Integer Sequences **18** (2015), 15p.

[24] Stefan Felsner, Sarah Kappes, Klemens Huemer, and David Orden, *Binary labelings for plane quadrangulations and their relatives*, Discr. Math. and Theor. Comp. Sci. (2010), 115–138.

[25] Stefan Felsner and Florian Zickfeld, *On the number of planar orientations with prescribed degrees*, Electron. J. Comb. **15** (2008).

[26] Philippe Flajolet and Robert Sedgewick, *Analytic Combinatorics*, Cambridge University Press, January 2009.

[27] Robert M. Gray, *Toeplitz and circulant matrices: A review*, Foundations and Trends in Communications and Information Theory **2** (2005), no. 3.

[28] Olle Hägström, *Finite Markov chains and algorithmic applications*, Cambridge University Press, Cambridge, 2002.

[29] Mahdieh Hasheminezhad and Brendan D. McKay, *Recursive generation of simple planar quadrangulations with vertices of degree 3 and 4*, Discussiones Mathematicae Graph Theory **30** (2010), no. 1, 123–136.

[30] Daniel Heldt, Kolja Knauer, and Torsten Ueckerdt, *On the bend-number of planar and outerplanar graphs*, Lecture Notes in Computer Science, LATIN **7256** (2012), 458–469.

[31] ______, *Edge-intersection graphs of grid paths: the bend-number*, Discrete Applied Mathematics **167** (2014), 144–162.

[32] Katherine Humphreys, *A history and a survey of lattice path enumeration*, Journal of Statistical Planning and Inference **140** (2010), 2237 – 2254.

[33] Mark Jerrum, *Counting, sampling and integrating: Algorithms and complexity*, Birkhäuser, Zürich, 2003.

[34] Mark Jerrum and Alistair Sinclair, *Approximate counting, uniform generation and rapidly mixing Markov chains*, Inf. Comput. **82** (1989), no. 1, 93–133.

[35] Kolja Knauer and Torsten Ueckerdt, *Three ways to cover a graph*, (submitted).

[36] Christian Krattenthaler, *The enumeration of lattice paths with respect to their number of turns*, Advances in Combinatorial Methods and Applications to Probability and Statistics (1997), 29–58.

[37] Steven Lalley, *Columbia university summer course: Stochastic interacting particle systems, problem set a: Monotone coupling*, 2007, (`http://galton.uchicago.edu/~lalley/Courses/Columbia/HWA.pdf`).

[38] David Levin, Yuval Peres, and Elizabeth Wilmer, *Markov chain and mixing times*, American Mathematical Society, 2009.

[39] Sarah Miracle, Dana Randall, Amanda P. Streib, and Prasad Tetali, *On sampling 3-orientations of planar triangulations*, Symposium on Discrete Algorithms, 2011, submitted., 2011.

[40] Gopal Mohanty, *Lattice path counting and applications*, Probability and mathematical statistics, Academic Press, 1979.

[41] Aleks Owczarek and Thomas Prellberg, *Enumeration of area-weighted Dyck paths with restricted height*, Australas. J. Combinat. **54** (2012), 13–18.

[42] Dana Randall and Prasad Tetali, *Analyzing Glauber dynamics by comparison of Markov chains*, Journal of Mathematical Physics **41** (1997), 1598–1615.

[43] Pierre Rosenstiehl, *Embedding in the plane with orientation constraints: The angle graph.*, Combinatorial mathematics, Proc. 3rd Int. Conf., New York/ NY (USA) 1985, Ann. N. Y. Acad. Sci. 555, 340 - 346, 1989.

[44] Pierre Rosenstiehl and Robert Endre Tarjan, *Rectilinear planar layouts and bipolar orientations of planar graphs.*, Discrete & Computational Geometry **1** (1986), 343–353.

[45] Neil Sloane, *The On-Line Encyclopedia of Integer Sequences*, (`http://oeis.org`).

[46] R P Stanley, *Enumerative combinatorics*, 2 ed., vol. 1, Wadsworth Publ. Co., Belmont, CA, USA, 1986.

[47] V. Strassen, *The existence of probability measures with given marginals*, The Annals of Mathematical Statistics **36** (1965), no. 2, pp. 423–439 (English).