Markov chain algorithms for Eulerian orientations and 3-colourings of 2-dimensional Cartesian grids

Johannes Fehrenbach and Ludger Rüschendorf
University of Freiburg

Abstract

In this paper we establish that the natural single point update Markov chain (also known as Glauber dynamics) for counting the number of Euler orientations of 2-dimensional Cartesian grids is rapidly mixing. This extends a result of Luby, Randall, and Sinclair (2001) who consider the case where orientations in the boundary are fixed. Similarly, we also obtain a rapid mixing result for the 3-colouring of rectangular Cartesian grids without fixing the boundaries. The proof uses path coupling and comparison to related Markov chains which allow additional transitions and which can be analysed directly.

Keywords: Markov chain algorithm, rapidly mixing, path coupling, Eulerian graphs, Cartesian grids, colouring problems

1 Introduction

Counting the number of Eulerian orientations of graphs is a problem with a variety of applications. For a 4-regular graph such as the (modified) 2-dimensional square lattice the number of Eulerian orientations, i.e. the number of distinct ways of orienting the edges in such a way that, at each vertex exactly two edges are directed in and two are directed out, is identical to the number of ice configurations in statistical physics (see Welsh (1990, 1993)). Mihail and Winkler (1996) proved that counting of Eulerian orientations in a general Eulerian graph G can be reduced to counting the perfect matchings in a related extended graph G' and thus they obtained a polynomial approximation algorithm for the general counting problem of however considerable high polynomial order.

In this paper we prove rapid mixing of the natural single point update Markov chain (also known as Glauber dynamics) for counting the Eulerian orientations of a rectangular grid when modified to be Eulerian. This leads to a considerable lower mixing rate compared to the general Mihail and Winkler algorithm. For rectangular grids with fixed boundary orientations a related rapid mixing result is obtained in Luby, Randall, and Sinclair (2001) for a Markov chain, which allows additional transitions. They relate this problem to the counting of routings. Randall and Tetali (1998) then established rapid mixing of the single point update Markov chain by a comparison argument.

In our paper we analyse in a first step for the Eulerian grid without fixed boundaries a related Markov chain, which admits additional transitions (*tower transitions*), as in Luby et al. (2001), by the path coupling method. Using the comparison method of Diaconis and Saloff-Coste (1993) this yields a rapid mixing result fur the natural one-step Markov chain.

There is a close relation between 3-colourings of rectangular grids and counting Eulerian orientations of the dual graph. A direct transformation of the Markov chain for the Euler orientations to the 3-colourings is however only possible when the boundary orientations and the boundary colourings are fixed. This connection was used in Luby et al. (2001) for the analysis of the 3-colouring problem. In section 4 of this paper we extend the one-point update Markov chain and a Markov chain with extended tower transitions to the framework of 3-colourings without fixed boundaries. By similar arguments and partially based on the results for the Eulerian orientations we establish rapid mixing of the 3-colouring Markov chains. For both problems the polynomial orders are of the same magnitude, only the constants differ. The extended chains are of the order n^3, where $n = a \cdot b$ is the magnitude of the grid. The one-point Glauber dynamics chains are of the order n^4a^2, where a is the maximal side length of the grid.

The rapid mixing property of the natural Markov chain for the counting of k-colourings has been studied in many recent papers. For the case of large number of colours (i.e. k approximatively $\geq 2\Delta$, Δ the maximal degree of the graph) very good mixing rate results have been obtained (see Jerrum (1995), Salas and Sokal (1997), Vigoda (2000), Bubley, Dyer, and Greenhill (1998)). For the case of small number of colours there remain several open problems. In particular the cases $k = 4, 5$ remain open in the Cartesian grid case.

The results in this paper are based on the dissertation of Fehrenbach (2003). After having finished this paper we got to know a recent preprint of Goldberg, Martin, and Paterson (2003) where the authors independently proved rapid mixing of the one-step chain and for a similarly extended Markov chain for the 3-colouring problem. The counting of Eulerian orientations is not considered in that paper. The rates in our paper are better than those in Goldberg et al.

(2003) which are of the order b^3a^6 for the extended chain and b^4a^9 for the Glauber dynamics (compared to b^3a^3 resp. b^4a^6 in our paper).

In our paper we use the path coupling method of Bubley and Dyer (1997). For an ergodic Markov chain $\mathcal{M}$ with transition matrix $P = (p_{ij})$ on a finite set Ω a (Markov-)coupling is a stochastic process $(X_t, Y_t)_{t\in\mathbb{N}}$ on $\Omega \times \Omega$ such that for all $x, y, z \in \Omega$ and $t \in \mathbb{N}$

$$P(X_{t+1} = x \mid X_t = y, Y_t = z) = p_{yx}$$
$$P(Y_{t+1} = x \mid X_t = y, Y_t = z) = p_{zx}$$

and $X_t = Y_t$ implies $X_{t+1} = Y_{t+1}$. Then by the coupling lemma

$$\left\|P^{X_t} - P^{Y_t}\right\| \leq P(X_t \neq Y_t), \tag{1.1}$$

where $\| \ \|$ is the variation norm.

Let $(X_t, Y_t)_{t\in\mathbb{N}}$ be a coupling and let $\delta : \Omega \times \Omega \to \mathbb{N}$ be a metric, such that for some $\beta \leq 1$ and all t

$$E(\delta(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t) = (x, y)) \leq \beta\delta(x, y). \tag{1.2}$$

Let $\tau(\varepsilon)$ be the mixing time of the Markov chain for the approximation error ε, i.e. $\tau(\varepsilon) = \sup_{x,y} \tau_{x,y}(\varepsilon)$ with $\tau_{x,y}(\varepsilon) = \inf\{t : \|P^{X_{t'}|X_0=x} - P^{Y_{t'}|Y_0=y}\| \leq \varepsilon$ for all $t' \geq t\}$ then

$$\tau(\varepsilon) \leq \frac{\log\left(\delta(\Omega)\varepsilon^{-1}\right)}{1-\beta} \qquad \text{if } \beta < 1. \tag{1.3}$$

If $\beta = 1$ and if for some $\alpha > 0$, and all $x, y \in \Omega$, $P\left(\delta(X_{t+1}, Y_{t+1}) \neq \delta(x, y) \mid (X_t, Y_t) = (x, y)\right) \geq \alpha$, $\forall t$, then

$$\tau(\varepsilon) \leq \frac{e\delta(\Omega)^2}{\alpha} \log \varepsilon^{-1}, \tag{1.4}$$

where $\delta(\Omega) = \max\{\delta(x, y) \mid x, y \in \Omega\}$ is the diameter of Ω (see Dyer and Greenhill (1998), Aldous (1983)).

The path coupling method is a technique which simplifies the construction of a coupling on all of $\Omega \times \Omega$ satisfying condition (1.2). It was introduced in Bubley and Dyer (1997). The following formulation is from Dyer and Greenhill (1998).

Let $S \subset \Omega \times \Omega$ be a set of transitions such that for all $x, y \in \Omega$ there exists a path $x = z_0, z_1, \ldots, z_r = y$ from x to y with transitions $(z_i, z_{i+1}) \in S$, $\forall i < r$. If $(x, y) \to (X', Y')$ is an $\mathcal{M}$-coupling for all $(x, y) \in S$, then an extension can be defined via the path in S for any state $(X_t, Y_t) = (x, y)$ in $\Omega \times \Omega$. One obtains thus a sequence $Z'_0, \ldots, Z'_r$ and a coupling $(X_{t+1}, Y_{t+1})_{t\in\mathbb{N}}$ on $\Omega \times \Omega$

with $X_{t+1} = Z'_0$ and $Y_{t+1} = Z'_r$. For a function $\Phi : S \to \mathbb{N}_0$ we define a metric $\delta(x, y) := \min \sum_{i=0}^{r-1} \Phi(z_i, z_{i+1})$, the minimum taken over all paths from x to y in S. If for some $\beta \le 1$

$$E(\delta(X', Y') \mid (X, Y) = (x, y)) \le \beta\delta(x, y) \text{ for all } (x, y) \in S, \tag{1.5}$$

then

$$E(\delta(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t) = (x, y)) \le \beta\delta(x, y) \text{ for all } (x, y) \in \Omega \times \Omega. \tag{1.6}$$

For the comparison of mixing times τ_1, τ_2 of two ergodic, reversible Markov chains $\mathcal{M}_1, \mathcal{M}_2$ with the same stationary distribution π and transition functions p_1, p_2 an effective method has been developed by Diaconis and Saloff-Coste (1993), see also Randall and Tetali (1998). Let $T_i = \{(x, y) \in \Omega \times \Omega \mid p_i(x, y) > 0\}$, $i = 1, 2$, be the graphs of $\mathcal{M}_1, \mathcal{M}_2$ and let $\Gamma = \{\gamma_{x,y}; (x, y) \in T_2\}$ be a set of *canonical* paths $\gamma_{x,y}$ in T_1 from x to y for any pair $(x, y) \in T_2$. For any $(w, z) \in T_1$ let $\Gamma(w, z) = \{(x, y) \in T_2; (w, z) \in \gamma_{x,y}\}$ be the set of all canonical paths in Γ which contain the edge (w, z). Finally, we denote by

$$A(\Gamma) := \max_{(w,z)\in T_1} \frac{1}{\pi(\omega)p_1(w, z)} \sum_{(x,y)\in\Gamma(w,z)} \pi(x)p_2(x, y)|\gamma_{x,y}|, \tag{1.7}$$

where $|\gamma_{x,y}|$ is the length of the path $\gamma_{x,y}$, the comparison measure. Then the following holds for all $\varepsilon \in \left(0, \frac{1}{2}\right)$:

$$\tau_1(\varepsilon) \le A(\Gamma) \frac{4\tau_2(\varepsilon)}{\log((2\varepsilon)^{-1})} \left(\log \hat{\pi}^{-1} + \log \varepsilon^{-1}\right) \tag{1.8}$$

where $\hat{\pi} = \min_{x\in\Omega} \pi(x)$.

This comparison result will be applied in sections 3 and 5 to determine bounds for the mixing time of the natural one-point improvement Markov chain (Glauber dynamics) by comparison with a chain which is simpler to analyse.

2 Eulerian orientations of Cartesian grids

An undirected, connected graph $G = (V, E)$ is called Euler graph if all vertices have even degree. A Eulerian orientation X of G is an orientation of the edges of G such that for each vertex $v \in V$ the set of edges directed towards v and the set of edges directed out of v have the same cardinality; i.e. with $E^+(v) := \{e = (v, w) \in X \mid w \in V\}$ and $E^-(v) := \{e = (w, v) \in X \mid w \in V\}$ holds: $|E^+(v)| = |E^-(v)|$ for all $v \in V$. Let $EO(G)$ denote the set of Eulerian orientations of G.

For a planar graph G let $F(G)$ denote the open domains generated by the embedding of G in $\mathbb{R}^2$. The inner domains are bounded while exactly one domain — the outer domain — is unbounded. For $\alpha \in F(G)$ and a Eulerian orientation let X^α denote the edges in the boundary of α directed according to $X \in EO(G)$. $\bar{e} := (w, v)$ denotes the inversion of the edge $e = (v, w)$; similarly for $C \subset V$, $\overline{C} := \{\bar{e}; e \in C\}$ denotes the inversion of the edges in C.

The following construction of a Markov chain $\mathcal{M}_0(G) = (X_t)_{t \in \mathbb{N}}$ on $EO(G)$ is quite natural.

Markov chain $\mathcal{M}_0(G)$ on $EO(G)$: To define the transition of $\mathcal{M}_0(G)$ from $X_t = x \in EO(G)$, we use two steps:

1) Let $\Lambda \in F(G)$ be a randomly sampled domain

2) If $\Lambda = \alpha$ and with $C := x^\alpha$ define

$$x' := \begin{cases} (x - C) \cup \overline{C} & \text{if } (x - C) \cup \overline{C} \in EO(G) \\ x & \text{else.} \end{cases} \tag{2.1}$$

Then define

$$X_{t+1} = \begin{cases} x' & \text{with probability } \frac{1}{2} \\ x & \text{else.} \end{cases} \tag{2.2}$$

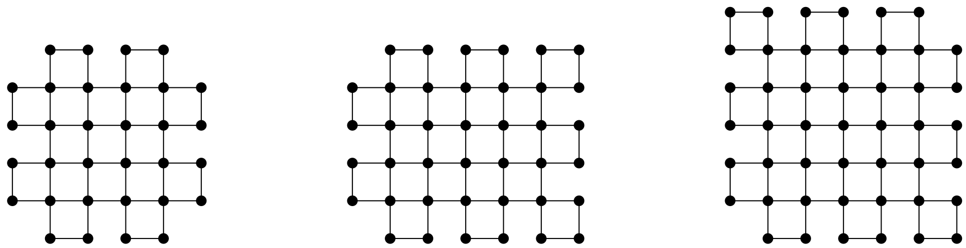
The corresponding transition matrix is denoted by P_0. The inversion of the edges in x^α for $\alpha \in F(G)$ is called α-transition.

Thus the Markov chain randomly chooses domains of the planar graph and inverts with probability $\frac{1}{2}$ the orientation of the boundary if possible. This Markov chain $\mathcal{M}_0(G)$ on $EO(G)$ is ergodic and the stationary distribution π of $\mathcal{M}_0(G)$ is the uniform distribution on $EO(G)$ (see Fehrenbach and Rüschendorf (2004)). It was proved in Fehrenbach and Rüschendorf (2004) that $\mathcal{M}_0(G)$ is rapidly mixing for planar triangular graphs. In this paper we consider rapid mixing for modified Cartesian grids. Let $(\overline{V}, \overline{E})$ denote a Cartesian $a \times b$ grid, where

$$\begin{aligned} \overline{V} &= \{0, \dots, a\} \times \{0, \dots, b\} \quad \text{and} \\ \overline{E} &= \{\{(i,j), (i+1,j)\} \mid 0 \le i < a, 0 \le j \le b\} \\ &\qquad \cup \{\{(i,j), (i,j+1)\} \mid 0 \le i \le a, 0 \le j < b\}. \end{aligned} \tag{2.3}$$

$(\overline{V}, \overline{E})$ is not a Eulerian graph but can be modified to a Eulerian graph by deleting edges at the outer domain, which lie at black inner domains as on a chess board (see Figure 2.1). Thus we obtain the *Euler $a \times b$ grid $G = (V, E)$*.

Luby et al. (2001) introduced a different modification of the $a \times b$ grid. They removed all edges in the boundary of the outer domain as well as the 4 vertices



The Eulerian 5×5 grid

The Eulerian 6×5 grid

The Eulerian 6×6 grid

Figure 2.1 Eulerian Cartesian grids

in the corner. The nodes at the boundary of the resulting grid $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ are of degree 1. The edges at the boundary are oriented clockwise in- and outwards and kept fixed as a kind of boundary condition. An orientation of the remaining edges is called Eulerian if $|E_X^-(v)| = |E_X^+(v)| = 2$ for all $v \in \widetilde{V}$ with degree $d_G(v) = 4$ (see Figure 2.2).

Figure 2.2 Grid with fixed boundary orientations

On these Eulerian orientations Luby et al. (2001) constructed a Markov chain with additional tower transitions which we also use for a modification $\mathcal{M}_0^*(G)$ of $\mathcal{M}_0(G)$ for the Euler $a \times b$ grid G. The mixing time of $\mathcal{M}_0^*(G)$ can be analysed directly by the path coupling method. By the comparison method we then obtain polynomial mixing rates for $\mathcal{M}_0(G)$.

Definition 2.1 (Tower transitions) *Let $G = (V, E)$ be a Euler $a \times b$ grid with $a, b > 1$ and $X \in EO(G)$. A subset $A = \{\alpha_1, \dots, \alpha_k\} \subset F(G)$ of domains of G is a tower of length k in X if*

a) A is a rectangle of side lengths 1 and k, where α_i, α_{i+1} are neighbour domains, $1 \le i < k$.

b) The symmetric difference $X^{\alpha_1} \oplus \dots \oplus X^{\alpha_k}$ is a directed circle.

c) X^{α_i} is only for $i = k$ a directed circle.

A tower A is maximal if there is no bigger tower $A' \supset A$, $A' \neq A$ (see Figure 2.3).

For a Euler $a \times b$ grid G and $X \in EO(G)$ any domain $\alpha \in F(G)$ which is a directed circle is a tower of length 1.

Lemma 2.2 *Let $A = \{\alpha_1, \ldots, \alpha_k\}$ be a tower of length k in $X \in EO(G)$, where G is a Euler $a \times b$ grid. If $\beta \in F(G) \cap A^C$ is an inner domain and neighbour to some $\alpha_i \in A$, then X^β is not a directed circle except when*

a) β is the neighbour of α_k opposite to α_{k-1} or

b) β is a neighbour of α_1.

Proof: Let $1 < i < k$, then the orientation of all edges at nodes at α_i are determined, as A is a tower and X a Eulerian orientation. If $\beta \notin A$ is a neighbour of α_i, then X^β is no directed circle (see Figure 2.3). For neighbours β of α_k and α_1, X^β can be a directed circle (see Figure 2.3). □

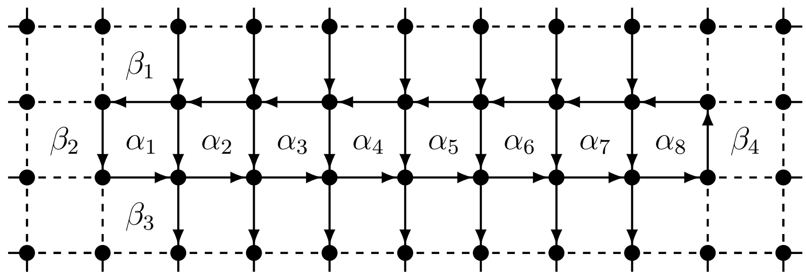


Figure 2.3 Tower transition $A = \{\alpha_1, \ldots, \alpha_8\}$, edges $\beta_1, \ldots, \beta_4$, and orientations of edges determined by A.

Using these towers we next introduce the Markov chain $\mathcal{M}_0^*(G)$ on the set of Eulerian orientations $EO(G)$ of a Euler $a \times b$ grid.

Definition 2.3 (Modified Markov chain $\mathcal{M}_0^*(G)$) *Let $G = (V, E)$ be the Euler $a \times b$ grid, $a, b > 1$. The Markov chain $\mathcal{M}_0^*(G) = (X_t)_{t \in \mathbb{N}}$ on $EO(G)$ is defined by the following transitions of $\mathcal{M}_0^*(G)$ from $X_t = x \in EO(G)$:*

1) Let $\Lambda \in F(G)$ be a randomly sampled inner domain.

2) If $\Lambda = \alpha_1$ and $\alpha_1, \ldots, \alpha_k$ is a tower of length k, then define $x' := (x - C) \cup \overline{C}$ where $C = x^{\alpha_1} \oplus \cdots \oplus x^{\alpha_k}$ and $X' := x$ else.

3) Define

$$X_{t+1} := \begin{cases} x' & \text{with probability } p \\ x & \text{else,} \end{cases} \tag{2.4}$$

where $p = \frac{1}{2}$ *if* $k = 1$ *and* $p = \frac{1}{4k}$ *if* $k > 1$. *Let* $P^*_0 = (p^*_0(\cdot,\cdot))$ *denote the transition matrix of* $\mathcal{M}^*_0(G)$.

Proposition 2.4 *The Markov chain* $\mathcal{M}^*_0(G)$ *is ergodic and the uniform distribution on* $EO(G)$ *is the stationary distribution* π *of* $\mathcal{M}^*_0(G)$.

Proof: Since any $\alpha \in F(G)$ such that x^α for some $x \in EO(G)$ is a directed circle is a tower of length 1, irreducibility of $\mathcal{M}^*_0(G)$ and aperiodicity of $\mathcal{M}^*_0(G)$ follow from the corresponding properties of $\mathcal{M}_0(G)$.

If $x, y \in EO(G)$ and $p^*_0(x,y) > 0$, $x \neq y$, then $C := x \oplus y = x^{\alpha_1} \oplus \cdots \oplus x^{\alpha_k}$ for some tower $A = \{\alpha_1, \dots, \alpha_k\}$. This implies however that also $B = \{\alpha_k, \dots, \alpha_1\}$ is a tower of length k and thus $p^*_0(y,x) = p^*_0(x,y)$; i.e. P^*_0 is symmetric. Together we obtain ergodicity of $\mathcal{M}^*_0(G)$. □

3 Rapid mixing of $\mathcal{M}^*_0(G)$ and $\mathcal{M}_0(G)$ for Cartesian grids

The mixing time τ^*_0 of the Markov chain $\mathcal{M}^*_0(G)$ on $EO(G)$ can be estimated by the path coupling method.

Theorem 3.1 (Rapid mixing of $\mathcal{M}^*_0(G)$) *Let* $G = (V, E)$ *be a Euler* $a \times b$ *grid with* $a, b > 1$ *and let* $\tau^*_0(\varepsilon)$ *denote the mixing time of* $\mathcal{M}^*_0(G)$. *Then for all* $\varepsilon \in (0, 1]$ *holds:*

$$\tau^*_0(\varepsilon) \leq \lceil 4en^3 \rceil \cdot \lceil \log \varepsilon^{-1} \rceil \tag{3.1}$$

where $n := |V|$ *and* e *is the Euler number.*

Proof: For the proof we apply the path coupling method with $\Omega = EO(G)$ the set of Eulerian orientations on G and

$$\mathcal{S} := \big\{(x,y) \in \Omega \times \Omega \mid \text{there exists an inner domain } \alpha \text{ such that } y = (x - x^\alpha) \cup \overline{x^\alpha}\big\}.$$

$\mathcal{S}$ thus is identical to the set of transitions of the Markov chain $\mathcal{M}_0(G)$. Therefore, by irreducibility of $\mathcal{M}_0(G)$ for any $x, y \in \Omega$ there exists a path $z_0, \dots, z_r$ form x to y with $z_0 = x$, $z_r = y$, and with transitions $(z_i, z_{i+1}) \in \mathcal{S}$. Define

$\Phi(x, y) = 1$ for $(x, y) \in \mathcal{S}$. Then $\delta(x, y)$ is the length of the shortest path from x to y in $\mathcal{S}$. We have to determine the coupling on $\mathcal{S}$.

Let $(x_1, y_1) \in \mathcal{S}$ with $y_1 = (x_1 - x_1^\alpha) \cup \overline{x_1^\alpha}$ for some inner domain $\alpha \in F(G)$. As the edges of x_1^α and y_1^α are a directed circle, it follows that x_1, y_1 can possibly differ only in towers which end in α for x_1 or y_1 or which begin in a neighbour domain of α. For these cases we next construct the couplings for the transitions of the Markov chain from $x_1 \to x_2$ and from $y_1 \to y_2$.

In step 1) of the construction of $\mathcal{M}_0^*(G)$ we choose for x_1 and y_1 the same inner domain β at random. Let $N(\alpha)$ denote the set of all neighbour domains of α. There are three cases to distinguish:

1. case: $\boldsymbol{\beta = \alpha}$

As x_1^α is a directed circle, the domain α is in x_1 and in y_1 a tower of length $k = 1$. By the second step of the construction of $\mathcal{M}_0^*(G)$, therefore, one obtains $x_1' = y_1$ and $y_1' = x_1$. Thus we define:

$$(X_2, Y_2) := \begin{cases} (x_1, x_1) & \text{with probability } \frac{1}{2} \\ (y_1, y_1) & \text{with probability } \frac{1}{2}. \end{cases} \tag{3.2}$$

Then $X_2 = Y_2$ with probability 1 and $\delta(X_2, Y_2) = \delta(x_1, y_1) - 1 = 0$.

2. case: $\boldsymbol{\beta \neq \alpha}$, $\boldsymbol{\beta \notin N(\alpha)}$

In β begins a tower A for the orientation x_1 exactly when a tower A' begins in β for y_1.

If $A = A' = \{\alpha_1, \ldots, \alpha_k\}$ is a tower of length k, then $C := x_1^{\alpha_1} \oplus \cdots \oplus x_1^{\alpha_k} = y_1^{\alpha_1} \oplus \cdots \oplus y_1^{\alpha_k}$. With $x_1' := (x_1 - C) \cup \overline{C}$, $y_1' := (y_1 - C) \cup \overline{C}$ we define according to step 3) of the definition of $\mathcal{M}_0^*(G)$

$$(X_2, Y_2) := \begin{cases} (x_1', y_1') & \text{with probability } p \\ (x_1, y_1) & \text{with probability } 1 - p, \end{cases} \tag{3.3}$$

where $p := \frac{1}{4k}$ for $k > 1$ and $p := \frac{1}{2}$ for $k = 1$. Then $\delta(X_2, Y_2) = \delta(X_1, Y_1) = 1$.

If $A \neq A'$, then α has to be a neighbour of some domain in A or has to be contained in A, as x_1, y_1 only differ in the orientations of the edges at α. As $\beta \notin N(\alpha)$ and $\beta \neq \alpha$ we obtain by Lemma 2.2 and Definition 2.1 that $\alpha = \alpha_k$ or α is the neighbour of α_k opposite to α_{k-1}. If the first case holds for x_1, then $A' = \{\alpha_1, \ldots, \alpha_{k-1}\}$, if the second case holds for x_1, then $A' = \{\alpha_1, \ldots, \alpha_{k+1}\}$ with $\alpha_{k+1} := \alpha$. One of both towers A, A' is by one element longer than the other. We assume w.l.o.g. that A is the longer tower and thus $\beta = \alpha_1$ and $\alpha_k = \alpha$. Further as $\beta \neq \alpha$ and $\beta \notin N(\alpha)$ we obtain $k > 2$. Define $C := x_1^{\alpha_1} \oplus \cdots \oplus x_1^{\alpha_k}$, $C' := y_1^{\alpha_1} \oplus \cdots \oplus y_1^{\alpha_{k-1}}$ and $x_1' := (x_1 - C) \cup \overline{C}$, $y_1' := (y_1 - C') \cup \overline{C'}$. As three edges from x_1^α are contained in C and the fourth edge is in C' we obtain $x_1' = y_1'$. We define the coupling by

$$(X_2, Y_2) = \begin{cases} (x_1', y_1') & \text{with probability } \frac{1}{4k} \\ (x_1, y_1') & \text{with probability } \frac{1}{4(k-1)} - \frac{1}{4k} \\ (x_1, y_1) & \text{with probability } \frac{1}{4(k-1)}. \end{cases} \tag{3.4}$$

Then

$$\delta(X_2, Y_2) = \begin{cases} \delta(X_1, Y_1) - 1 & \text{with probability } \frac{1}{4k} \\ \delta(X_1, Y_1) + k - 1 & \text{with probability } \frac{1}{4(k-1)} - \frac{1}{4k} \\ \delta(X_1, Y_1) & \text{with probability } \frac{1}{4(k-1)}. \end{cases}$$

Thus we obtain for the expected difference of the distances of this coupling

$$\begin{aligned} \Delta &:= E\left(\delta(X_2, Y_2) \mid (X_1, Y_1) = (x_1, y_1)\right) - \delta(x_1, y_1) \\ &= (-1)\frac{1}{4k} + (k-1)\left(\frac{1}{4(k-1)} - \frac{1}{4k}\right) = 0. \end{aligned}$$

3. case: $\boldsymbol{\beta \in N(\alpha)}$

In this case not the same tower can begin in β for x_1 and for y_1, as the edge between α and β has different orientations in x_1 and in y_1 (cf. Lemma 2.2). We discuss the possible subcases one after the other.

3a) A tower of length $\boldsymbol{k > 2}$ begins for $\boldsymbol{x_1}$ in $\boldsymbol{\beta}$:

Then no tower begins in β for the orientation y_1. We define the coupling

$$(X_2, Y_2) := \begin{cases} (x_1', y_1) & \text{with probability } \frac{1}{4k} \\ (x_1, y_1) & \text{with probability } 1 - \frac{1}{4k}, \end{cases} \tag{3.5}$$

with x_1' corresponding to step 2 of the definition of $\mathcal{M}_0^*(G)$. Then we obtain for the expected change of the distance $\Delta(\beta)$ in this case $\Delta(\beta) = k \cdot \frac{1}{4k} = \frac{1}{4}$.

3b) $\boldsymbol{\beta}$ is a tower of length 1 for $\boldsymbol{x_1}$:

In this case in β begins the tower $A = \{\alpha_1, \alpha_2\}$ with $\alpha_1 := \beta$, $\alpha_2 = \alpha$ for y_1. This results in $x_1' = y_1'$ and we define

$$(X_2, Y_2) = \begin{cases} (x_1', y_1') & \text{with probability } \frac{1}{8} \\ (x_1', y_1) & \text{with probability } \frac{1}{2} - \frac{1}{8} \\ (x_1, y_1) & \text{with probability } \frac{1}{2}. \end{cases} \tag{3.6}$$

As a result we obtain $\Delta(\beta) = (-1)\frac{1}{8} + 1(\frac{1}{2} - \frac{1}{8}) = \frac{1}{4}$.

3c) A tower A of length 2 begins for β:
Let $A = \{\beta, \beta'\}$. If $\beta' = \alpha$, then for y_1 a tower of length 1 begins in β. This is identical to case 3b) with the roles of x_1, y_1 switched. If $\beta' \neq \alpha$ then for y_1 no tower begins in β and we define the coupling as in case 3a). As consequence we obtain for the expected change of distance $\Delta(\beta)$ in this case $\Delta(\beta) = \frac{1}{4}$.

3d) No tower begins for x_1 in β:
Then two situations are possible. Either a tower begins in β for y_1. This leads to the same situation as in case 3a) with switched roles of x_1, y_1. Or no tower begins in β for the orientation y_1. Then we define $(X_2, Y_2) := (x_1, y_1)$ and, therefore, $\delta(X_2, Y_2) = \delta(x_1, y_1) = 1$.

From these situations we obtain together: If β is a neighbour of α then $\Delta(\beta) \leq \frac{1}{4}$.

As a result of cases 1)–3) finally we obtain for the expected difference of the distance

$$\begin{aligned}\Delta &= E\left(\delta(X_2, Y_2) - \delta(X_1, Y_1) \mid (X_1, Y_1) = (x_1, y_1)\right)\\ &\leq \frac{1}{|F(G)| - 1}\left((-1) + |N(\alpha)| \cdot \frac{1}{4}\right) \leq 0.\end{aligned}$$

Further we obtain for the variation of the Markov chain the inequality

$$P\left(\delta(X_2, Y_2) \neq \delta(x_1, y_1) \mid (X_1, Y_1) = (x_1, y_1)\right) \geq \alpha := \frac{1}{|F(G)| - 1} \tag{3.7}$$

The proof of (3.7) is by induction on the length k of the shortest path form x_1 to y_1 in $\mathcal{S}$. The case $k = 1$ follows from the first case. For the induction step consider a path of length $k + 1$, $x_0 x_1 \ldots x_k = y_1$, $d' = \delta(x_0, y_1) = k + 1$, and the corresponding transitions $X_0, X_1, \ldots, X_k = Y_1$. We denote $\delta' = \delta(X_0, Y_1)$, $\delta = \delta(X_1, Y_1)$, $d' = \delta(x_0, y_1)$, $d = \delta(x_1, y_1)$. Then conditionally on $X_0 = x_0$, $X_1 = x_1$, $Y_1 = y_1$ we obtain:

$$\begin{aligned}P(\delta' \neq d') &= P(\delta' > d') + P(\delta' < d')\\ &\geq P(\delta' > d' \mid \delta > d)P(\delta > d) + P(\delta' < d' \mid \delta \leq d)P(\delta \leq d)\end{aligned}$$

By case 1, $P(\delta' < d' \mid \delta \leq d) \geq \alpha$ and from part 3) of the definition of the Markov chain $\mathcal{M}_0^*(G)$ holds $P(\delta' > d' \mid \delta > d) \geq 1 - p \geq \frac{1}{2} > \alpha$. Therefore, $P(\delta' \neq d') \geq \alpha$.

With $D := \max_{x,y \in \Omega} \delta(x, y) \leq 2(|F(G)| - 1)$, $\beta := 1$ and $\alpha := \frac{1}{|F(G)|-1}$ we obtain from the path coupling estimate in (1.4)

$$\begin{aligned}\tau_0^*(\varepsilon) &\leq \left\lceil \frac{e(2(|F(G)| - 1))^2}{(|F(G)| - 1)^{-1}} \right\rceil \log \varepsilon^{-1}\\ &\leq \lceil 4e(|F(G)| - 1)^3 \rceil \lceil \log \varepsilon^{-1} \rceil. \end{aligned} \tag{3.8}$$

We have $m = |E| \leq 2n$ and from the Euler polyeder formula $n + \ell - m = 2$ for planar graphs, where $\ell = |F(G)|$ we obtain $\ell < n$, which leads to the final estimate

$$\tau_0^*(\varepsilon) \leq \lceil 4en^3 \rceil \lceil \log \varepsilon^{-1} \rceil. \tag{3.9}$$

□

In the next step we apply the comparison technique in (1.8) to obtain a bound for the mixing time τ_0 of the single point update Markov chain $\mathcal{M}_0(G)$.

Theorem 3.2 (Rapid mixing of $\mathcal{M}_0(G)$) *Let $G = (V, E)$ be a Eulerian $a \times b$ grid with $a \geq b > 1$ and let τ_0 denote the mixing time of the single point update chain $\mathcal{M}_0(G)$. Then for all $\varepsilon \in (0, \frac{1}{2})$*

$$\tau_0(\varepsilon) \leq 44n^3a^2(2n + \log \varepsilon^{-1}), \tag{3.10}$$

where $n = |V|$.

Proof: In order to compare the mixing time $\tau_0(\varepsilon)$ with $\tau_0^*(\varepsilon)$ for $\mathcal{M}_0^*(G)$ we have to construct a set Γ of canonical paths $\gamma_{x,y}$ in $\mathcal{M}_0(G)$ for each transition (x, y) in $\mathcal{M}_0^*(G)$, $\Gamma = \{\gamma_{x,y} \mid (x, y) \text{ a transition in } \mathcal{M}_0^*(G)\}$.

Let $x, y \in EO(G)$ and let $A = \{\alpha_1, \dots, \alpha_k\}$ be a tower in x such that $y = (x - C) \cup \overline{C}$, where $C = x^{\alpha_1} \oplus \cdots \oplus x^{\alpha_k}$. $\gamma_{x,y}$ is constructed by induction on the length k of the tower. If $k = 1$ then the transition of $\mathcal{M}_0^*(G)$ is already a transition of $\mathcal{M}_0(G)$ and $\gamma_{x,y}$ is defined by this transition. If $k > 1$, then x^{α_k} is a directed circle. The Eulerian orientation $x' = (x - x^{\alpha_k}) \cup \overline{x}^{\alpha_k}$ differs from y in the orientation of the edges $x^{\alpha_1} \oplus \cdots \oplus x^{\alpha_{k-1}}$. Thus a transition from x' to y in $\mathcal{M}_0^*(G)$ is defined via the tower $A' = \{\alpha_1, \dots, \alpha_{k-1}\}$ of length $k - 1$. By induction hypothesis the path $\gamma_{x'y} = (x_i')_{0 \leq i < k}$ is already defined. We define then the canonical path from x to y by $\gamma_{xy} = (x_i)_{0 \leq i \leq k}$, where $x_0 := x$ and $x_{i+1} := x_i'$ for $0 \leq i < k$.

We denote for transitions (w, z) of $\mathcal{M}_0(G)$ by $\Gamma(w, z)$ the set of all canonical paths which contain (w, z), i.e. $\Gamma(w, z) := \{(x, y) \in \Omega^2 \mid (w, z) \in \gamma_{x,y}\}$. We have to estimate the comparison measure $A(\Gamma)$ in (1.7). The maximal length of a tower in G is at most $\max\{a, b\}$. Further, an inner domain $w \in F(G)$ such that (w, z) is a transition in $\mathcal{M}_0(G)$, is contained at most in $a + b - 2$ further towers, i.e. $|\Gamma(w, z)| \leq a + b - 1$. This implies

$$\begin{aligned} A(\Gamma) = & \max_{\substack{(w,z)\in\Omega^2 \\ p_0(w,z)>0}} \frac{1}{\pi(w)p_0(w, z)} \sum_{(x,y)\in T(w,z)} \pi(x)p_0^*(x, y)|\gamma_{x,y}| \\ \leq & \max_{\substack{(w,z)\in\Omega^2 \\ p_0(w,z)>0}} \sum_{(x,y)\in\Gamma(w,z)} \frac{p_0^*(x, y)}{p_0(w, z)} \max\{a, b\} \end{aligned}$$

$$
\begin{aligned}
&\le \max_{\substack{(w,z)\in\Gamma^2\\ p_0(w,z)>0}} |\Gamma(w,z)| \max\{a,b\} \\
&\le (a+b-1)\max\{a,b\} \ \le\ 2\max\{a,b\}^2 \qquad (3.11)
\end{aligned}
$$

Further, $|EO(G)| \le 2^m$ with $m = |E|$ and $m \le 2n$ and, therefore, $\widehat{\pi} = \min_{x\in\Omega} \pi(x) \ge 4^{-n}$. Thus from (1.8) and Theorem 3.1 we obtain

$$
\begin{aligned}
\tau_{_0}(\varepsilon) &\le \frac{4\log(\widehat{\pi}\varepsilon)^{-1}}{\log(2\varepsilon)^{-1}} a^2 \tau_{_0}^*(\varepsilon) \\
&\le \lceil 16en^3 \rceil a^2 \frac{\lceil \log\varepsilon \rceil}{\log 2\varepsilon}(2n+\log\varepsilon^{-1}) \\
&\le 44n^3a^2(2n+\log\varepsilon^{-1}) \text{ for all } \varepsilon \in \left(0,\tfrac{1}{2}\right). \qquad \square
\end{aligned}
$$

Thus for $a = b$ the polynomial order of the Glauber dynamics Markov chain is n^5 while the extended chain is of the order n^3. Since the comparison argument in Theorem 3.2 is very rough it should be possible to prove that the order of $\tau_{_0}(\varepsilon)$ is close to n^3, maybe with some logarithmic factor.

4 3-colourings and Eulerian orientations

The problem of counting Eulerian orientations of an $a \times b$ grid G is related to the problem of counting the 3-colourings of the vertices of G.

A *k-colouring* (of vertices) of a graph $G = (V,E)$ is a mapping $\Phi : V \to \{0,\dots,k-1\}$ such that $\Phi(v) \ne \Phi(w)$ for all $e = \{v,w\} \in E$. Let $F^k(G)$ denote the set of all k-colourings of G.

Jerrum (1995) introduced a simple Markov chain $\mathcal{M}_C(G)$ for counting the colourings of low degree graphs. This chain chooses at random one node $v \in V$ and a colour $c \in \{0,\dots,k-1\}$ and replaces the colour of v by c if this leads to an admissible colouring. Our aim is to prove rapid mixing of $\mathcal{M}_C(G)$ for the 3-colourings of $a \times b$ grids.

Let for a graph $G = (V,E)$, $G^* = (V^*, E^*)$ denote the dual graph with node set V^* the set of domains $F(G)$ and where two domains α, α' are connected in E^* if α, α' have a common boundary in G. For an $a \times b$ grid $G = (V,E)$ the dual graph is Eulerian. The boundary of each inner domain has 4 edges, while at the outer domain there are $2(a+b)$ edges. Baxter (1982) established the following connection between 3-colourings of G and Euler orientations of G^*.

If $\Phi \in F^3(G)$, then one obtains a Eulerian orientation $X = X_\Phi \in EO(G^*)$ defining for any $e = \{v,w\} \in E$ the orientation of $e^* = \{\alpha,\beta\}$ as follows: If v, w are vertical neighbours i.e. $v = (i,j)$ and $w = (i,j+1)$, then e^* is a horizontal edge in G^* with α the left and β the right end node. Define:

$$
\begin{array}{rll}
 & (\alpha,\beta) \in X & \text{if and only if } \Phi(v) = \Phi(w) - 1 \mod 3 \\
\text{and} & (\beta,\alpha) \in X & \text{if and only if } \Phi(v) = \Phi(w) + 1 \mod 3.
\end{array} \qquad (4.1)
$$

Figure 4.1

If v, w are horizontal neighbours, i.e. $v = (i, j)$, $w = (i + 1, j)$, then e^* is a vertical edge in G^* with lower end note α and upper end node β. Then we define

$$\begin{array}{lll} & (\alpha, \beta) \in X & \text{iff } \Phi(v) = \Phi(w) + 1 \mod 3 \\ \text{and} & (\beta, \alpha) \in X & \text{iff } \Phi(v) = \Phi(w) - 1 \mod 3. \end{array} \tag{4.2}$$

Figure 4.2

(4.1), (4.2) define a Eulerian orientation. For a node $\alpha \in V^*$ not corresponding to the outer domain of G there are two different colourings. Both lead to two edges in α and two edges out of α (see Figure 4.3).

Figure 4.3

This implies that also the node corresponding to the outer domain of G has the same number of in- and outgoing nodes.

If, conversely, X is a Eulerian orientation of G^*, then a 3-colouring cannot be uniquely defined. The colour of the first node is arbitrary while the colours of the outer nodes then are uniquely defined by the orientation. Thus we obtain

$$|F^3(G)| = 3|EO(G^*)|. \tag{4.3}$$

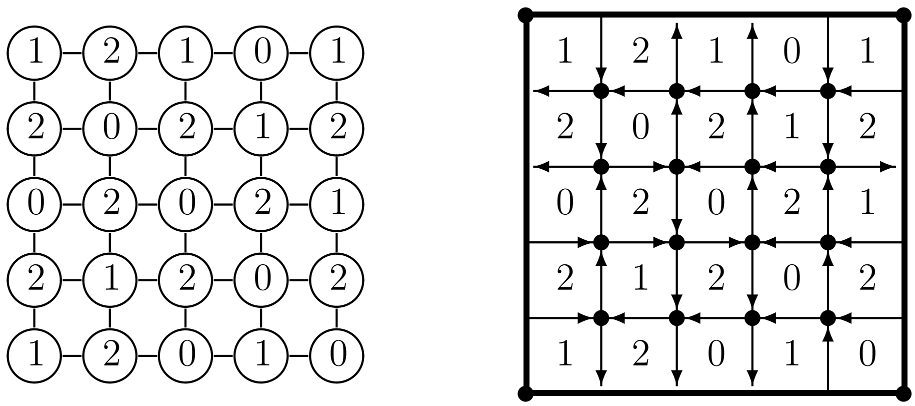


Figure 4.4 3-colourings of a 5×5 grid and Eulerian orientation of the dual grid.

This connection was used in (Luby et al. 2001) to transfer Eulerian orientations of the $a \times b$ grid $\widetilde{G}$ to 3-colourings of the dual grid in the case of fixed boundaries. Their consideration of orientations with fixed boundaries leads to uniquely defined Markov chains on the set of 3-colourings with fixed colouring of the boundaries. In the following we construct directly the Markov chain $\mathcal{M}_C(G)$ on the set of all colourings $F^3(G)$ on the $a \times b$ grid G with uniform distribution as the stationary distribution of $\mathcal{M}_C(G)$.

As seen above $EO(G)$ is not bijective to $F^3(G^*)$ and thus $\mathcal{M}_0(G^*)$ does not determine uniquely a Markov chain on $F^3(G)$. To define an analogous Markov chain on the 3-colourings we observe that for $\Phi \in F^3(G)$ a node v can change the colour only if all neighbours of v have the same colour. In this case the edges in G^* at the dual node v^* are a directed circle in the orientation of G^* corresponding to Φ. Thus we obtain the following Markov chain $\mathcal{M}_C(G)$ mentioned already in the beginning of this section.

Definition 4.1 (Markov chain $\boldsymbol{\mathcal{M}_C(G)}$ on $\boldsymbol{F^3(G)}$) *Let $G = (V, E)$ be an $a \times b$ grid in the plane. To define the transitions of $\mathcal{M}_C(G) = (\Phi_t)_{t \in \mathbb{N}}$ from $\Phi_t = \Phi \in F^3(G)$ we use three steps:*

1) *Let N_t be a randomly sampled node $v \in V$ and let C_t be a randomly sampled colour $c \in \{0, 1, 2\}$.*

2) *If $N_t = v$ and $C_t = c$, then define $\Phi' : V \to \{0, 1, 2\}$ by*

$$\Phi'(w) = \begin{cases} c & \text{if } w = v \\ \Phi(w) & \text{if } w \neq v. \end{cases}$$

3) *Define*

$$\Phi_{t+1} = \begin{cases} \Phi' & \text{if } \Phi' \in F^3(G) \\ \Phi & \text{else.} \end{cases} \tag{4.4}$$

The corresponding transition matrix of $\mathcal{M}_C(G)$ *is denoted by* $P_C = (p_C(\cdot,\cdot))$.

Proposition 4.2 *For any* $a \times b$ *grid* G *with* $a, b > 1$, $\mathcal{M}_C(G)$ *is ergodic. The stationary distribution of* $\mathcal{M}_C(G)$ *is the uniform distribution on* $F^3(G)$.

Proof: Aperiodicity and reversibility of $\mathcal{M}_C(G)$ are consequences of the definition. We have to establish that $\mathcal{M}_C(G)$ is irreducible. As described above any $X \in EO(G^*)$ leads to a 3-colouring Φ_X of G by choosing for the *first* node v^0 of G any colour from $\{0, 1, 2\}$. The other colours are then determined by the orientation X. A transition of $\mathcal{M}_0(G^*)$, therefore, corresponds to a transition of $\mathcal{M}_C(G)$, if the in the first step of the definition of $\mathcal{M}_0(G^*)$ sampled domain of G^* does not correspond to v^0. Thus the graph of $\mathcal{M}_C(G)$ decomposes into at most 3 components, each having a fixed colour of v^0. But any of these components also contains 2-colourings of G, so that also node v^0 can change the colour. This implies that the graph of $\mathcal{M}_C(G)$ is irreducible.

Further any $\Phi, \Psi \in F^3(G)$ are connected by a path in $\mathcal{M}_C(G)$ of length $\leq 2n$. To see this we consider two cases. If $\Phi(v^0) = \Psi(v^0)$, then this follows from the corresponding statement for $\mathcal{M}_0(G)$. If $\Phi(v^0) \neq \Psi(v^0)$, say $\Phi(v^0) = 0$, $\Psi(v^0) = 1$, then starting from Φ a 2-colouring Θ of G with $\Theta(v^0) = 0$ can be reached in $\leq n$ steps. There is a positive transition from Θ to

$$\Theta'(w) := \begin{cases} \Psi(v^0), & w = v^0, \\ \Phi(w), & w \neq v^0, \end{cases} \quad \text{for } w \in V.$$

Θ is connected to Ψ by a path of length $\leq n$ as $\Psi(v^0) = \Theta'(v^0)$. Combining these paths we obtain a path from Φ to Ψ of length $\leq 2n$. □

In order to analyse $\mathcal{M}_C(G)$ we introduce similar to section 2 a modification $\mathcal{M}^*_C(G)$ of $\mathcal{M}_C(G)$ which allows additional tower transitions. This is possible as the dual graph G^* resembles the Eulerian $a \times b$ grid in section 2.

Definition 4.3 *Let* $G = (V, E)$ *be a Cartesian* $a \times b$ *grid and* $\Phi \in F^3(G)$. *A subset* $A = \{v_1, \dots, v_k\} \subset V$ *is called tower of length* k *in* Φ *if*

1) For all $i < k$, $e_i = \{v_i, v_{i+1}\} \in E$,

2) either $\Phi(v_i) = \Phi(v_{i+1}) + 1 \mod 3$, $\quad 0 \leq i < k$,
or $\quad \Phi(v_i) = \Phi(v_{i+1}) - 1 \mod 3$, $\quad 0 \leq i < k$,

3) if $k > 1$ *and* $\{w, v_1\} \in E$, $w \neq v_2$, *then* $\Phi(w) \neq \Phi(v_2)$,

4) if $\{v_k, w\} \in E$, $\{v_k, w'\} \in E$, *then* $\Phi(w) = \Phi(w')$.

A tower A of Φ possibly contains nodes of the boundary i.e. of degree < 4. These nodes correspond to domains in G^* which have only 2 or 3 inner edges. The analogue of Lemma 2.2 holds true.

Lemma 4.4 *If $\Phi \in F^3(G)$ and $A = \{v_1, \ldots, v_k\}$ is a tower of length k in Φ, then there exist at most 4 nodes $\{w_1, \ldots, w_4\} \subset V \setminus A$ which are neighbours of nodes in A and whose neighbours have the same colour in Φ, i.e. for all $i \leq 4$, $w', w'' \in N(w_i)$ holds $\Phi(w') = \Phi(w'')$. The w_i are either some neighbours of $v_1, w_i \neq v_2$ or the neighbour of v_k opposite to v_{k-1}.*

Proof: For $k = 1$ the statement is obvious. For $k > 1$ it follows by induction that point 3) in Definition 4.3 is true not only for v_1 but also for all $v_i \in A$ with $i < k$, as the nodes in $N(v_i)\setminus A$ are also neighbours of the nodes $w \in N(v_{i-1})\setminus A$ which already satisfy $\Phi(w) \neq \Phi(v_i)$. Thus the nodes in $N(v_i) \setminus A$ are coloured like v_{i-1}, which themselves by 2) in Definition 4.3 have a different colour as v_{i+1}. Therefore, only the nodes in $N(v_1) \setminus \{v_2\}$ or the neighbour of v_k opposite to v_{k-1} possibly have the same colour as their neighbours in A. □

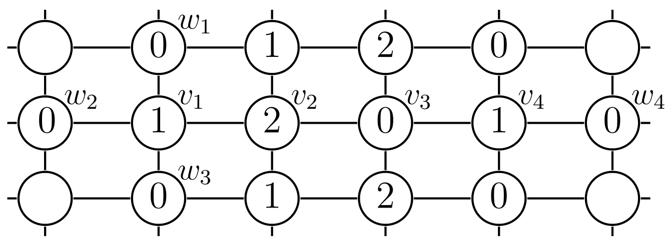


Figure 4.5 Tower A and the neighbours w_i with identical colour and some further colours determined by the tower.

Definition 4.5 (Modified Markov chain $\mathcal{M}_C^*(G)$) *Let $G = (V, E)$ be a Cartesian $a \times b$ grid. The Markov chain $\mathcal{M}_C^*(G) = (\Phi_t)_{t\in\mathbb{N}}$ on $F^3(G)$ is defined by the following transitions from $\Phi_t = \Phi \in F^3(G)$:*

1) *Let N_t be a randomly sampled node and let C_t be a randomly sampled colour in $\{0, 1, 2\}$.*

2) *If $N_t = v_1$, $C_t = c$ and $A = \{v_1, \ldots, v_k\}$ is a tower of length k in Φ beginning in v_1, then define $\Phi' : V \to \{0, 1, 2\}$ by*

$$\Phi'(w) := \begin{cases} \Phi(w) + \varphi \mod 3 & \text{if } w \in A \\ \Phi(w) & \text{if } w \notin A, \end{cases}$$

where $\varphi := c - \Phi(v_1) \mod 3$.

3) *If $\Phi' \in F^3(G)$ then define*

$$\Phi_{t+1} := \begin{cases} \Phi' & \text{with probability } p \\ \Phi & \text{with probability } 1 - p, \end{cases}$$

where $p = 1$ if $k = 1$, $p = \frac{1}{2k}$ if $k > 1$.

If $\Phi' \notin F^3(G)$ then define $\Phi_{t+1} = \Phi$. The transition matrix of $\mathcal{M}_C^*$ is denoted by $P_C^* = (p_C^*(\cdot,\cdot))$.

Proposition 4.6 *For an $a \times b$ grid $G = (V, E)$ the Markov chain $\mathcal{M}_C^*(G)$ is ergodic. The stationary distribution π of $\mathcal{M}_C^*(G)$ is the uniform distribution on $F^3(G)$.*

Proof: Each transition of $\mathcal{M}_C(G)$ is also a transition of $\mathcal{M}_C^*(G)$. Therefore, $\mathcal{M}_C^*(G)$ is irreducible. As for any $\Phi \in F^3(G)$, $p_C^*(\Phi,\Phi) > 0$ $\mathcal{M}_C^*(G)$ is aperiodic. If a transition from Φ to Ψ is given in $\mathcal{M}_C^*(G)$ by a tower $A = \{v_1, \ldots, v_k\}$ and with colour $c = \Psi(v_1)$, then A is also a tower for the colouring Ψ with begin in v_k and end in v_1 and with colour $c' = \Phi(v_k)$. Further, for these transitions holds

$$p_C(\Phi,\Psi) = p_C(\Psi,\Phi) = \begin{cases} \frac{1}{3n} & \text{if } k = 1 \\ \frac{1}{6nk} & \text{if } k > 1. \end{cases}$$

Thus $\mathcal{M}_C^*(G)$ is also reversible and, therefore, ergodic. □

5 Rapid mixing property of the 3-colourings Markov chains

In analogy to the results in section 3 we next establish the rapid mixing property of the Markov chains $\mathcal{M}_C^*(G)$ and then by the comparison argument the rapid mixing property for $\mathcal{M}_C(G)$.

Theorem 5.1 (Mixing time of the modified chain $\mathcal{M}_C^*(G)$) *For an $a \times b$ grid $G = (V, E)$ with $a, b > 1$ the Markov chain $\mathcal{M}_C^*(G)$ on the set of 3-colourings $F^3(G)$ is rapidly mixing. For the mixing time $\tau_C^*(\varepsilon)$ holds the bound*

$$\tau_C^*(\varepsilon) \leq 18n^3 \lceil \log \varepsilon^{-1} \rceil \tag{5.1}$$

for $\varepsilon \in (0, 1]$ and $n := |V| = a \cdot b$.

Proof: We apply the path coupling method with $\Omega = F^3(G)$ and

$$\mathcal{S} = \left\{ (\Phi, \Psi\} \in \Omega^2 \,\middle|\, |\{v \in V \mid \Phi(v) \neq \Psi(v)\}| = 1 \right\}.$$

By irreducibility of $\mathcal{M}_C(G)$ there exists for any $\Phi, \Psi \in \Omega$ a path from Φ to Ψ in $\mathcal{S}$. The metric δ on Ω is defined as the length of the shortest path in $\mathcal{S}$. We have to define the coupling on $\mathcal{S}$.

Let $(\Phi_1, \Psi_1) \in \mathcal{S}$ and let $v \in V$ be the node with $\Phi_1(v) \neq \Psi_1(v)$, w.l.o.g. $\Phi_1(v) = 0$, $\Psi_1(v) = 1$. Then for all $v' \in N(v)$, $\Phi_1(v') = \Psi_1(v') = 2$. Let in step 2 of the definition of $\mathcal{M}_C^*(G)$, $N_t = w \in V$ and $C_t = c \in \{0, 1, 2\}$. Then we determine the coupling (Φ_2, Ψ_2) as follows:

1. case: $\boldsymbol{w = v}$
As $\Phi_1(v') = \Psi_1(v') = 2$ for all $v' \in N(v)$, there begins in v a tower in Φ_1 and in Ψ_1 of length 1. By the second step of the construction of $\mathcal{M}_C^*(G)$ then $\Phi_1' = \Psi_1$ if $c = 1$ and $\Psi_1' = \Phi_1$ if $c = 0$. If $c = 2$, then Φ_1' and Ψ_2' are not in $F^3(G)$. Therefore, we define

$$(\Phi_2, \Psi_2) := \begin{cases} (\Phi_1, \Phi_1) & \text{if } c = 0 \\ (\Psi_1, \Psi_1) & \text{if } c = 1 \\ (\Phi_1, \Psi_1) & \text{if } c = 2. \end{cases} \tag{5.2}$$

Then $\Phi_2 = \Psi_2$ with probability $\frac{2}{3}$ and thus conditionally on $w = v$

$$E(\delta(\Phi_2, \Psi_2) - \delta(\Phi_1, \Psi_1) \mid v) = -\frac{2}{3}. \tag{5.3}$$

2. case: $\boldsymbol{w \neq v}$ and $\boldsymbol{w \notin N(v)}$
As $w \notin N(v)$ we have $\Phi_1(w') = \Psi_1(w')$ for all $w' \in N(w)$. Therefore, a tower A begins in w in Φ_1 if and only if a tower A' begins in w in Ψ_1.

If $\mathbf{A} = \mathbf{A}' = \{w_1, \ldots, w_k\}$, then Φ_1' and Ψ_1' only differ in the colour of v independent of the choice of c. Then in case $c = \Phi_1(w_2) = \Psi_1(w_2)$ we define

$$(\Phi_2, \Psi_2) := \begin{cases} (\Phi_1', \Psi_1') & \text{with probability } \frac{1}{2k} \\ (\Phi_1, \Psi_1) & \text{with probability } 1 - \frac{1}{2k} \end{cases}$$

and $(\Phi_2, \Psi_2) := (\Phi_1, \Psi_1)$ iff $c \neq \Phi_1(w_2)$. In both cases holds

$$\delta(\Phi_2, \Psi_2) = \delta(\Phi_1, \Psi_1) = 1 \tag{5.4}$$

If $\mathbf{A} \neq \mathbf{A}'$, then by Lemma 4.4 either A or A' ends in node v as both begin in the same node. Let w.l.o.g. $A = \{w_1, \ldots, w_k\}$ end in v; $w_k = v$, $w_1 = w$. Then $A' = \{w_1, \ldots, w_{k-1}\}$ and as $w \notin N(v)$ we have $k > 2$. By step 2 of the definition of $\mathcal{M}_C^*(G)$ $\Phi_1' = \Psi_1'$ if $c = \Phi_1(w_2)$. We define the coupling

$$(\Phi_2, \Psi_2) := \begin{cases} (\Phi_1', \Psi_1') & \text{with probability } \frac{1}{2k} \\ (\Phi_1, \Psi_1') & \text{with probability } \frac{1}{2(k-1)} - \frac{1}{2k} \\ (\Phi_1, \Psi_1) & \text{with probability } 1 - \frac{1}{2(k-1)}. \end{cases}$$

Then

$$\delta(\Phi_2,\Psi_2) := \begin{cases} \delta(\Phi_1,\Psi_1)-1 & \text{with probability } \frac{1}{2k} \\ \delta(\Phi_1,\Psi_1)+k-1 & \text{with probability } \frac{1}{2(k-1)}-\frac{1}{2k} \\ \delta(\Phi_1,\Psi_1) & \text{with probability } 1-\frac{1}{2(k-1)}. \end{cases}$$

The expected difference of the distance conditionally on w is

$$\begin{aligned} \Delta(w) &:= E(\delta(\Phi_2,\Psi_2)-\delta(\Phi_1,\Psi_1) \mid w) \\ &= \frac{1}{3}\left((-1)\frac{1}{2k}+(k-1)\left(\frac{1}{2(k-1)}-\frac{1}{2k}\right)\right)=0. \end{aligned} \tag{5.5}$$

In case $c \neq \Phi_1(w_2)$ we define $(\Phi_2,\Psi_2) := (\Phi_1,\Psi_1)$ and obtain

$$\delta(\Phi_2,\Psi_2)=\delta(\Phi_1,\Psi_1)=1. \tag{5.6}$$

3. case: $w \in N(v)$
In this case not the same tower can begin in w for Φ_1 as for Ψ_1 (cf. Lemma 4.4). We discuss the possible subcases.

3a) For Φ_1 a tower A of length $k > 2$ begins in w:
If $A = \{w_1,\dots,w_k\}$, then for Ψ_1 no tower begins in w. In case $c = \Phi_1(w_2)$ define

$$(\Phi_2,\Psi_2) = \begin{cases} (\Phi_1',\Psi_1) & \text{with probability } \frac{1}{2k} \\ (\Phi_1,\Psi_1) & \text{with probability } 1-\frac{1}{2k}; \end{cases}$$

in case $c \neq \Phi_1(w_2)$ define $(\Phi_2,\Psi_2) := (\Phi_1,\Psi_1)$. As a result

$$\Delta(w) = \frac{2}{3}\cdot 0 + \frac{1}{3}\cdot k \cdot \frac{1}{2k} = \frac{1}{6} \tag{5.7}$$

3b) For Φ_1 a tower of length 1 begins in w:
The neighbours of w have all the same colour for Φ_1. As $\Phi_1(w) \neq \Psi_1(w)$ there begins the tower $A := \{w, v\}$ for Ψ_1 in w and by step 2 of the definition of $\mathcal{M}_C^*(G)$ we obtain $\Phi_1' = \Psi_1'$. The coupling is defined by

$$(\Phi_2,\Psi_2) = \begin{cases} (\Phi_1',\Psi_1') & \text{with probability } \frac{1}{4} \\ (\Phi_1',\Psi_1) & \text{with probability } 1-\frac{1}{4}, \end{cases}$$

if $c = \Psi_1(v)$. Then we obtain

$$\Delta(w) = \frac{1}{3}\left((-1)\frac{1}{8}+1\cdot\left(\frac{1}{2}-\frac{1}{8}\right)\right) = \frac{1}{12}. \tag{5.8}$$

If $c \neq \Psi_1(v)$, then $(\Phi_2,\Psi_2) := (\Phi_1,\Psi_1)$ and thus $\delta(\Phi_2,\Psi_2) = 1$.

3c) For Φ_1 a tower A of length 2 begins in w:
If $v \in A$, then w is for Ψ_1 a tower of length 1. This is identical to case 3b) with the roles of Φ_1, Ψ_1 switched. If $v \notin A$, then no tower for Ψ_1 begins in w and we define the coupling as in case 3a). Thus we obtain in this case

$$\Delta(w) \leq \frac{1}{6}. \tag{5.9}$$

3d) For Φ_1 no tower begins in w:
If for Ψ_1 a tower begins in w, then we have a situation as in case 3a) with roles of Φ_1, Ψ_1 switched. If for Ψ_1 no tower begins in w, then define $(\Phi_2, \Psi_2) := (\Phi_1, \Psi_1)$ and $\delta(\Phi_2, \Psi_2) = \delta(\Phi_1, \Psi_1) = 1$.

From these cases we obtain: If w is a neighbour of v, then

$$\Delta(w) \leq \frac{1}{6} \tag{5.10}$$

As a result of the cases 1), 2), 3) we finally obtain for the expected difference of the distance conditioned under $(\Phi_1, \Psi_1) \in \mathcal{S}$:

$$\begin{aligned}\Delta &= E(\delta(X_2, Y_2) - \delta(\Phi_1, \Psi_1)) \\ &\leq \frac{1}{n} \cdot \left(-\frac{2}{3} + |N(v)| \cdot \frac{1}{6} \right) \leq 0.\end{aligned} \tag{5.11}$$

Similarly, to the argument in the proof of (3.7) we obtain from case 1 by an induction argument on the length k that

$$P\left(\delta(\Phi_2, \Psi_2) \neq \delta(\Phi_1, \Psi_1)\right) \geq \alpha := \frac{2}{3n}. \tag{5.12}$$

With $D = \max_{\Phi,\Psi\in\Omega} \delta(\Phi, \Psi) \leq 2n$ and with $\beta := 1$ we obtain from (1.4) and the argument in the proof of (3.9) the following estimate of the mixing time τ_C^*:

$$\tau_C^*(\varepsilon) \leq \left\lceil \frac{e(2n)^2}{\left(\frac{2}{3n}\right)^{-1}} \right\rceil \cdot \lceil \log \varepsilon^{-1} \rceil \leq 18n^3 \cdot \lceil \log \varepsilon^{-1} \rceil. \tag{5.13}$$

□

Finally the comparison technique in (1.7) yields a bound for the mixing time τ_C of the one-step colouring Markov chain $\mathcal{M}_C(G)$.

Theorem 5.2 (Mixing time of the one-step colouring chain $\mathcal{M}_C(G)$) *Let $G=(V,E)$ be an $a\times b$ grid with $a\geq b>1$ and let τ_C denote the mixing time of the one-step Markov chain $\mathcal{M}_C(G)$ for the 3-colourings $F^3(G)$. Then for all $\varepsilon\in(0,\frac{1}{2})$ and with $n=|V|=a\times b$ holds:*

$$\tau_C(\varepsilon)\leq 144\cdot n^3a^2(n\log 3+\log\varepsilon^{-1}). \tag{5.14}$$

Proof: We compare τ_C with the mixing time τ_C^* of $\mathcal{M}_C^*(G)$. Let $\Phi,\Psi\in F^3(G)$ and let $A=\{v_1,\ldots,v_k\}$ be a tower for Φ such that for $v\in V$:

$$\Psi(v)=\begin{cases}\Phi(v)+\varphi \mod 3 & \text{if } v\in A\\ \Phi(v) & \text{if } v\notin A,\end{cases}$$

where $\varphi:=\Psi(v_1)-\Phi(v_1) \mod 3$.

This transition can be replaced by k transitions in $\mathcal{M}_C(G)$, which are constructed by induction on k. The case $k=1$ is obvious. If $k>1$ then we define

$$\Phi'(v):=\begin{cases}\Psi(v) & \text{if } v=v_k\\ \Phi(v) & \text{if } v\neq v_k.\end{cases}$$

By definition the neighbours of v_k all have the colour $\phi(v_{k-1})$. As $\Psi\in F^3(G)$, we have $\Psi(v_k)\neq\phi(v_{k-1})$. Thus $\Phi'\in F^3(G)$ and Φ' and Ψ differ by the tower $A'=\{v_1,\ldots,v_{k-1}\}$. By the assumption of the induction the path $\gamma_{\Phi',\Psi}=(\Phi'_i)_{0\leq i<k}$ is already constructed in $\mathcal{M}_C(G)$. We define $\gamma_{\Phi,\Psi}=(\Phi_i)_{0\leq i\leq k}$ by $\Phi_0:=\Phi$ and $\Phi_{i+1}:=\Phi'_i$ for $0\leq i<k$. This defines the set Γ of canonical paths for all $(\Phi,\Psi)\in\Omega^2$, $\Omega=F^3(G)$. For any transition $(\Theta,\Upsilon)\in\mathcal{S}$ of $\mathcal{M}_0(G)$ define

$$\Gamma(\Theta,\Upsilon):=\{(\Phi,\Psi)\in\Omega^2\mid(\Theta,\Upsilon)\in\mathcal{S}_{\Phi\Psi}\}. \tag{5.15}$$

We have to estimate the comparison measure $A(\Gamma)$ in (1.7). The maximal length of a tower in G is at most $\max\{a,b\}$. Further an inner node w which is for Θ a tower of length 1 is at most contained in $a+b-2$ further towers and thus $|\Gamma(\Theta,\Upsilon)|\leq a+b-1$ for all transitions $(\Theta,\Upsilon)$of $\mathcal{M}_C(G)$. This implies

$$\begin{aligned}A(\Gamma) &= \max_{\substack{(\Theta,\Upsilon)\in\Omega^2\\ p_C(\Phi,\Upsilon)>0}}\frac{1}{\pi(\Phi)p_C(\Theta,\Upsilon)}\sum_{(\Phi,\Psi)\in\Gamma(\Theta,\Upsilon)}\pi(\Phi)p_C^*(\Phi,\Psi)|\gamma_{\Phi,\Psi}|\\ &= \max_{\substack{(\Theta,\Upsilon)\in\Omega^2\\ p_C(\Phi,\Upsilon)>0}}\sum_{(\Phi,\Psi)\in\Gamma(\Theta,\Upsilon)}\frac{p_C^*(\Phi,\Psi)}{p_C(\Theta,\Upsilon)}\max\{a,b\}\\ &\leq \max_{\substack{(\Theta,\Upsilon)\in\Omega^2\\ p_C(\Theta,\Gamma)>0}}|\Gamma(\Theta,\Upsilon)|\max\{a,b\}\\ &\leq (a+b-1)\max\{a,b\}\leq 2a^2.\end{aligned}$$

Further $\widehat{\pi} = \min_{\Phi \in F^3(G)} \pi(\Phi) \geq 3^{-n}$ and thus from the comparison theorem (1.8) and Theorem 5.1 we obtain

$$\begin{aligned}\tau_C(\varepsilon) &\leq \frac{8\log(\widehat{\pi}\varepsilon)^{-1}}{\log(2\varepsilon)^{-1}} a^2 \tau_C^*(\varepsilon) \\ &\leq 144 n^3 a^2 \frac{[\log \varepsilon]}{\log 2\varepsilon}(n \log 3 + \log \varepsilon^{-1}) \\ &\leq 144 n^3 a^2 (n \log 3 + \log \varepsilon^{-1})\end{aligned}$$

for all $\varepsilon \in (0, \frac{1}{2})$. □

References

Aldous, D. (1983). *Random walks on finite groups and rapidly mixing Markov chains*, Volume 986 of *Lecture Notes in Mathematics*, pp. 243–297. Springer.

Baxter, R. (1982). *Exactly solved models in statistical mechanics*. Academic Press.

Bubley, R. and M. Dyer (1997). Path coupling: a technique for proving rapid mixing in Markov chains. In *Proceedings of the 38th Annual IEEE Symposium on foundations of Computer Science (FOCS)*, pp. 223–231. IEEE Computer Society Press.

Bubley, R., M. Dyer, and C. Greenhill (1998). Beating the 2δ bound for approximately counting colourings: a computer-assisted proof of rapid mixing. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, Volume 9, pp. 355–363.

Diaconis, P. and L. Saloff-Coste (1993). Comparison theorems for reversible Markov chains. *Annals of Applied Probability 3*, 696–730.

Dyer, M. and C. Greenhill (1998). A more rapidly mixing Markov chain for graph colorings. *Random Structure and Algorithms 13*, 285–317.

Dyer, M. and C. Greenhill (1999). Random walks on combinatorial objects. In J. Lamb and D. Preece (Eds.), *Surveys in Combinatorics*, pp. 101–136. Cambridge Universtiy Press.

Fehrenbach, J. (2003). *Design und Analyse stochastischer Algorithmen auf kombinatorischen Strukturen*. PhD thesis, University of Freiburg.

Fehrenbach, J. and L. Rüschendorf (2004). A Markov chain algorithm for Eulerian orientations of planar triangular graphs. To appear in: Proceedings of the Third Colloquium on Mathematics and Computer Sciences, Algorithms, Trees, Combinatorics, and Probabilities, Vienna.

Goldberg, L., R. Martin, and M. Paterson (2003). Random sampling of 3-colourings in $\mathbb{Z}^2$. Preprint, University of Warwick.

Jerrum, M. (1995). A very simple algorithm for estimating the number of k-colourings of a low-degree graph. *Random Structures and Algorithms 7*, 157–165.

Luby, M., D. Randall, and A. Sinclair (2001). Markov chain algorithms for planar lattice structures. *SIAM Journal on Computing 31*, 167–192.

Mihail, M. and P. Winkler (1996). On the number of Eulerian orientations of a graph. *Algorithmica 16*, 402–414.

Randall, D. and P. Tetali (1998). *Analyzing Glauber dynamics by comparison of Markov chains*, Volume 1380 of *Lecture Notes in Computer Science*, pp. 292–304. Springer.

Salas, J. and A. Sokal (1997). Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *Journal of Statistical Physics 86*, 551–579.

Vigoda, E. (2000). Improved bounds for sampling colourings. *Journal of Mathematical Physics 41*, 1555–1569.

Welsh, D. (1990). The computational complexity of some classical problems from statistical physics. In G. Grimmett and J. Hammersley (Eds.), *Disorder in Physical Systems*, pp. 307–321. Oxford University Press.

Welsh, D. (1993). *Complexity: Knots, Colourings, and Counting*, Volume 186 of *LMS Lecture Note Series*. Cambridge University Press.

Johannes Fehrenbach
Department of Mathematics
University of Freiburg
Eckerstr. 1
79104 Freiburg
Germany

Ludger Rüschendorf
Department of Mathematics
University of Freiburg
Eckerstr. 1
79104 Freiburg
Germany

ruschen@stochastik.uni-freiburg.de