Snapshots of modern mathematics
from Oberwolfach

№ 2/2016

Random sampling of domino and lozenge tilings

Éric Fusy [1]

A *grid region* is (roughly speaking) a collection of "elementary cells" (squares, for example, or triangles) in the plane. One can "tile" these grid regions by arranging the cells in pairs. In this snapshot we review different strategies to generate random tilings of large grid regions in the plane. This makes it possible to observe the behaviour of large random tilings, in particular the occurrence of boundary phenomena that have been the subject of intensive recent research.

1 Introduction

Tilings of regular grids have been intensively studied due to their rich combinatorial properties and the intriguing and challenging macroscopic phenomena which occur in random tilings of certain regions when the size gets large.

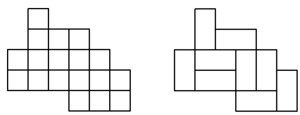

Figure 1: A square grid region and a domino tiling.

In this snapshot, we review methods for the random generation of tilings, focusing on two classical types of tilings: domino tilings and lozenge tilings.

[1] Éric Fusy is partly supported by the ANR grant "Cartaplus" 12-JS02-001-01 and the ANR grant "EGOS" 12-JS02-002-01.

First, we review in Section 3 a random generator for domino tilings of the Aztec Diamond, which relies on a combinatorial identity for these tilings discovered by Elkies, Kuperberg, Larsen, and Propp [1] and is described in the companion snapshot [7], which we assume the reader has read. However, this combinatorial identity is specific to the Aztec Diamond, and it seems difficult to adapt the strategy to general regions. We then discuss in Section 4 a technique based on Markov chains that can be applied in a wide variety of cases. Along the way, we also discuss running time issues (which is crucial to sample random tilings of large size).

2 Domino and lozenge tilings

A *square grid region* is a finite set of *elementary cells* of the square grid $\mathbb{Z}^2$, each elementary cell being of the form $[x, x+1] \times [y, y+1]$ for some $(x, y) \in \mathbb{Z}^2$ (see Figure 1). Similarly, a *triangular grid region* is a finite set of elementary cells of the (regular) triangular grid, each cell being an equilateral triangle of side length 1 pointing either left or right, see Figure 2. Two cells of a (square or triangular) grid region R are called *adjacent* if they share an edge. We will only consider *connected* regions here, that is, regions such that for each pair of cells c_1, c_2 of the region R there exists a sequence of successively adjacent cells of R with c_1 as first element and c_2 as last element.

A *domino tiling* of a square grid region R is a way of grouping the cells of R into adjacent pairs; equally, *lozenge tiling* of a triangular grid region R is a way of grouping the cells of R into adjacent pairs, see Figure 1 and Figure 2 for examples. Note that it is not always possible to tile a region R. (A clear obstacle is when R has an odd number of cells.) We denote by $\mathcal{T}(R)$ the set of (domino or lozenge, depending on the grid) tilings of the grid region R. Whenever there are any possible (domino or lozenge) tilings of a grid region R, we are interested in designing procedures that output a random (domino or lozenge) tiling, each tiling having the same probability $1/\mathrm{Card}(\mathcal{T}(R))$ of being chosen.[2] We call this *sampling uniformly at random from* the set of tilings $\mathcal{T}(R)$. A motivation is that this allows to observe the qualitative picture of a large random tiling, especially the boundary behaviour (see Section 5).

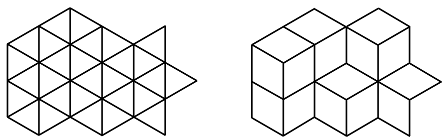

Figure 2: A triangular grid region and a lozenge tiling of the region.

[2] Recall that the *cardinality* $\mathrm{Card}(S)$ of a finite set S is the number of elements it contains.

A first naive approach for the random sampling from $\mathcal{T}(R)$ is to *list* all the tilings in $\mathcal{T}(R)$, and then pick up a tiling from the list at random. However, the set $\mathcal{T}(R)$ is usually very large. (We have seen in the companion snapshot [7] that for the Aztec Diamond of order n, $\mathcal{T}(R)$ has cardinality $2^{n(n+1)/2}$, which is already 3 602 879 7018 963 968 for $n = 10$, far beyond computer possibilities!) In the algorithms we will describe, rather than choosing among a list of already constructed tilings, the random tiling is generated "on the fly", that is, the tiling can be seen as progressively obtained from a sequence of random choices. In this way we can obtain much more efficient algorithms, whose running time has a *polynomial* (rather than *exponential*) dependency on the size parameter n.[3]

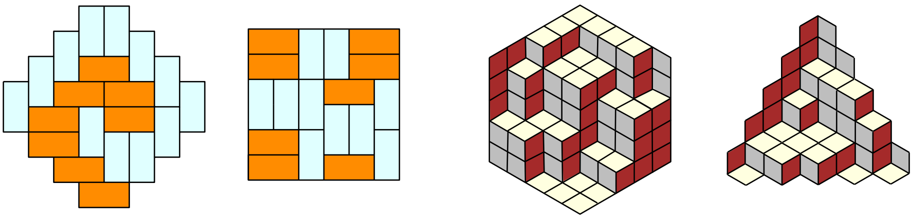

Figure 3: Examples of tilings of regions with a regular shape, from left to right: (1) domino tiling of the Aztec Diamond (of order 4), (2) domino tiling of the 6×6 square, (3) lozenge tiling of the $5 \times 5 \times 5$ hexagon (can be seen as a heap of cubes stacked in a $5 \times 5 \times 5$ three dimensional box), (4) lozenge tiling of the $5 \times 5 \times 5$ equilateral triangle region. The colourings of the lozenges give 3D representations of the tilings.

[3] In computer science, it is always desirable to know about the *time complexity* of an algorithm, that is, how fast its running time τ_n grows when the size n of the input data gets larger. For example, quadratic dependency on the size parameter n means that the running time τ_n grows roughly like n^2 when n is large (this is written as $\tau_n = \mathcal{O}(n^2)$), and, more in general, polynomial dependency means that it grows approximately like n^k for some constant k (in symbols, $\tau_n = \mathcal{O}(n^k)$). Polynomial running times are always preferable to exponential running times since the exponential function grows faster than every polynomial.

3 A random sampler for the Aztec Diamond

The *Aztec Diamond of order n* is the square grid region R_n made of the cells that are entirely included in $\{|x| + |y| \leq n + 1\}$. (An Aztec Diamond of size 4 is displayed in the leftmost picture in Figure 3.) As shown in [1] and also explained in the companion snapshot [7], $\mathcal{T}(R_n)$ has cardinality $2^{n(n+1)/2}$, and there is a combinatorial way to prove it using alternating sign matrices. [4]

We denote by $\mathcal{A}_n$ the set of alternating sign matrices of size $n \times n$. In [7] we have seen that the set of domino tilings of the Aztec Diamond $\mathcal{T}(R_n)$ can be identified with the set of pairs of "compatible" matrices,

$$\{(A, B) : \; A \in \mathcal{A}_n, \; B \in \mathcal{A}_{n+1}, \; A \sim B\},$$

where the tilde $A \sim B$ denotes some compatibility relation between the matrices A and B, which is given explicitly in [7]. For an alternating sign matrix A we denote by $N_-(A)$ and $N_+(A)$ the number of -1 and $+1$ entries in A, respectively. We have seen in the snapshot [7] that for $A \in \mathcal{A}_n$, there are $2^{N_+(A)}$ matrices $B \in \mathcal{A}_{n+1}$ such that B is compatible to A (in symbols, $A \sim B$), and for $B \in \mathcal{A}_{n+1}$ there are $2^{N_-(B)}$ matrices $A \in \mathcal{A}_n$ such that $A \sim B$. This allows to define (recursively) the following random sampler for $\mathcal{T}(R_n)$:

- If $n = 1$, output either  or , each with probability 1/2,
- If $n > 1$,
  1. first, output a random tiling T in $\mathcal{T}(R_{n-1})$ by calling the random sampler at order $n-1$. Let $(A, B) \in \mathcal{A}_{n-1} \times \mathcal{A}_n$ be the associated pair of matrices,
  2. then pick uniformly at random a matrix $A' \in \mathcal{A}_{n+1}$ such that $B \sim A'$ (there are $2^{N_+(B)}$ such matrices, and as it turns out, one can choose A' by tossing an unbiased coin $N_+(B)$ times),
  3. return the tiling in $\mathcal{T}(R_n)$ associated to (B, A').

One can prove by induction on n that the distribution at size n is indeed the uniform distribution, that is, every tiling has the same chance to get picked. The sampler can equivalently be formulated using height functions since these are closely related to the alternating sign matrices, as explained in the snapshot [7]. Figure 4 shows an example of an execution formulated with the height function. (It can be checked that the question marks in the 5th picture correspond indeed to the $+1$ entries of B) It is easy to turn the recursive procedure into an iterative procedure where a domino tiling is "grown" from order 1 to order n, and moreover the procedure can equivalently be formulated using so-called "domino shuffling" operations [1], see also the nice survey article [2], which gives animated executions of the domino shuffling iterative random sampler.

---

[4] Recall from the companion snapshot [7] that an $n \times n$-matrix M is called an *alternating sign matrix* if its entries are in $\{+1, 0, -1\}$ and if in each row or column, the sequence of non-zero entries alternates between 1 and -1, starting with 1 and ending with 1.

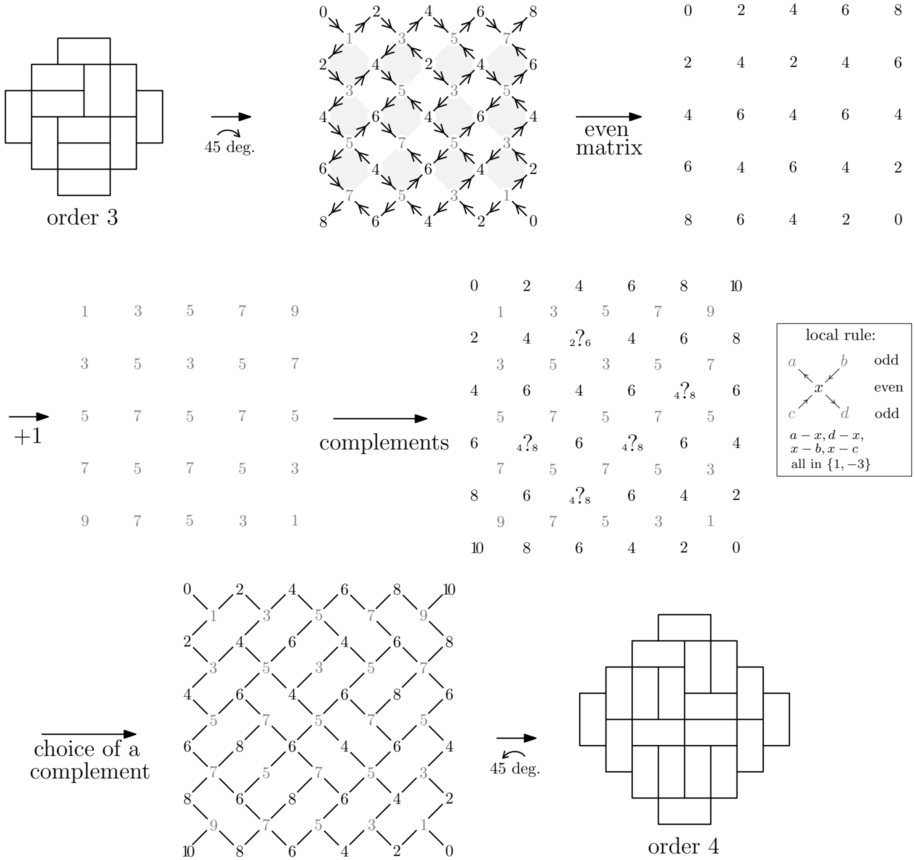

Figure 4: From a random domino tiling of order 3 to a random domino tiling of order 4. The tiling of order 3 is rotated clockwise by 45 degrees, and the height function is computed, giving two matrices gathering the rows of even and odd rank, respectively. The odd matrix is discarded, and the coefficients of the even matrix are increased by 1 (making it an odd matrix). The next drawing shows all the even matrices that are compatible with the obtained odd matrix. (Question marks correspond to the coefficients with two possible values.) Choosing such a compatible matrix uniformly at random (by tossing a coin for each question mark entry), we obtain a random domino tiling of order 4.

As for the time complexity, it can be shown by induction that the running time τ_n of a call to the sampler at order n satisfies $\tau_n = \tau_{n-1} + \sigma_n$, where σ_n consists of operations (such as computing the height function and choosing a compatible matrix) that each require a traversal of the Aztec regions R_{n-1} or R_n, whose number of cells is quadratic in n. Therefore, τ_n is cubic in n, that is, $\tau_n = O(n^3)$.

To conclude, the Aztec Diamond is a very nice kind of region in the sense that its tilings can be set in correspondence with a simpler set where the random generation can be efficiently performed. But such a correspondence can in general not be established for arbitrary regions, [5] and it is therefore desirable to have a uniform random sampling strategy that works for any region. This is described in the next section.

4 Random samplers based on local moves

We now describe a random sampling strategy that works for any grid region R with non-empty $\mathcal{T}(R)$. We just require that R is *hole-free*, meaning that for any connected "necklace" of cells of R, the entire area inside the necklace is covered by cells belonging to R. The strategy relies on *local moves*, which we describe for domino tilings. (The approach for lozenge tilings is very similar.)

Given a domino tiling of a square grid region R, a *flip* consists in turning a pair of adjacent horizontal dominoes into a pair of vertical dominoes (inside the surrounding 2×2 square), and a *flop* is the reverse operation, see the upper-left part of Figure 5. If the grid region R is hole-free, then it can be shown that the set $\mathcal{T}(R)$ of possible tilings is *connected* under the flip-flop relations [8], that is, for any tilings T_1, T_2 in $\mathcal{T}(R)$, one can move from T_1 to T_2 by a sequence of flip/flop moves (see the right part of Figure 5).

We can now describe a simple process based on local moves: [6] start from an arbitrary tiling in $\mathcal{T}(R)$ and at each step do the following operations:

1. choose uniformly at random a 2×2 square S of cells of R,
2. if the square S is made of two adjacent horizontal (respectively vertical) dominoes, then with probability $1/2$ perform a flip (respectively a flop) inside S. In all the other situations, stay idle.

---

[5] The number L_n of lozenge tilings of the $n \times n \times n$ hexagon also satisfies a simple formula: $L_n = \prod_{1 \le i,j,k \le n} \frac{i+j+k-1}{i+j+k-2}$. There is a combinatorial proof of this formula [4] (based this time on "15-puzzle principles") giving rise to a uniform random sampler, also of time complexity $O(n^3)$.

[6] For lozenge tilings, one uses the local moves in the lower-left part of Figure 5, which in the 3D representation correspond to adding/removing a unit cube.

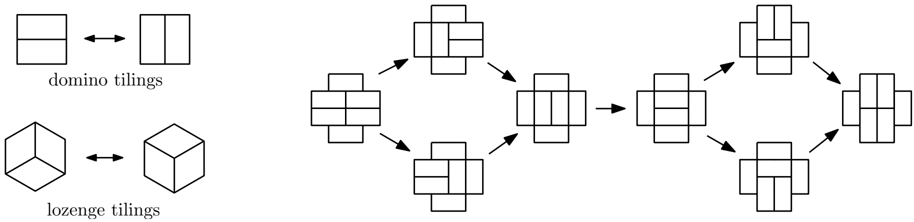

Figure 5: Left: local moves for domino and lozenge tilings. Right: the local move relations for domino tilings of the Aztec Diamond of order 2 (note that the set $\mathcal{T}(R)$ is connected under these relations).

In that process, it is clear that what is decided at each step only depends on the current state and not on the past history; such a process is called a *Markov chain*. This Markov chain satisfies a *symmetry property* (the probability of moving in one step from one tiling to another tiling is the same as the probability of returning to the original tiling in the next step), which implies that the uniform distribution is stable for this process, that is, if the probability distribution of the current state is uniform, then the probability distribution of the next state is also uniform. Combined with a few other properties that need to be checked (in particular the fact that the flip/flop relations make $\mathcal{T}(R)$ a connected set), it implies that, starting from any fixed tiling, the distribution on $\mathcal{T}(R)$ approaches the uniform distribution as the number of executed steps gets large.

In general, there is a "threshold" time, called the *mixing time* t_{mix}, such that after t_{mix} steps the distribution gets quickly close to the uniform distribution. In our case, we typically have a family $\{R_n : n \in \mathbb{N}\}$ of regions indexed by a side-length parameter $n \in \mathbb{N}$ (such as the coloured regions shown in Figure 3), so that for each n we have a mixing time $t_{\text{mix}}(n)$, which quantifies how long the Markov chain has to be run at size n to get close enough to the uniform distribution. It is strongly believed that for any natural family $\{R_n : n \in \mathbb{N}\}$ of regions indexed by a side-length parameter n, $t_{\text{mix}}(n)$ should grow like $n^4 \log(n)$ (with log denoting the logarithm function); however, rigorous bounds of the form $t_{\text{mix}}(n) = O(n^4 \log(n))$ are still unproved. [7] It is therefore more desirable to have a procedure with a guaranteed uniform distribution; the seminal papers [5, 6] give such a strategy called "coupling from the past", where

[7] A bound $t_{\text{mix}}(n) = O(n^4 \log(n))$ is proved in [9] for a slightly non-local modification of the Markov chain on lozenge tilings, where at each step a pile of cubes (instead of a single cube) might be added or deleted.

instead of running the Markov chain forward, it is run “from the past”, with the benefit that it automatically detects when to stop in order to output a tiling under the uniform distribution. The algorithm is stopped once a certain “coalescence” is reached, and the coalescence time is again expected grow like $n^4 \log(n)$, which is not far from the time complexity $O(n^3)$ of the random sampler of Section 3 (specific to the Aztec Diamond).

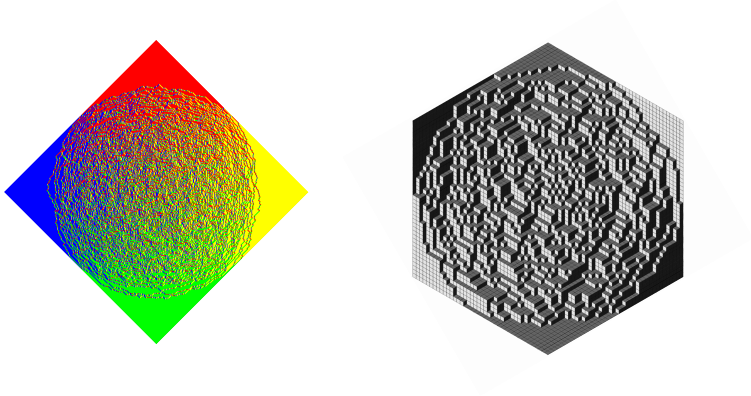

Figure 6: Left: a random domino tiling of the Aztec Diamond of order 512. Right: a random lozenge tiling of the $40 \times 40 \times 40$ hexagon.

5 Boundary phenomena

A motivation for random generation of tilings is to observe and conjecture the behaviour of random structures when the size gets large. Figure 6 shows uniformly random tilings of large size, for the Aztec Diamond (domino tiling) and the hexagon (lozenge tiling). One observes two rules: inside the circle tangent to the outer frame, the tiling shows randomness (disorder), outside of that circle the dominoes/lozenges are aligned and the tiling is totally ordered (*frozen regions*). Such phenomena tend to occur when the boundaries of the region are such that it is difficult to locally change the tiling close to the boundary. This is the case for the Aztec Diamond and the hexagon. For lozenge tilings of the $n \times n \times n$ hexagon one can observe a simple occurrence of this phenomenon: if we try to have a corner of the hexagon bordering on two lozenges (instead of one), then the constraint propagates along the two sides meeting at this corner and determines the tiling around the border of the hexagon. Such a tiling thus reduces to a tiling of the $(n-1) \times n \times n$ hexagon, which has much fewer tilings, so that such a configuration at a corner of the hexagon is very unlikely. In contrast, for the square grid region and equilateral triangle grid region (second and forth drawing in Figure 3), no macroscopic frozen region can be observed. This so-called *arctic circle phenomenon*, and more generally the study of phase transitions in random tilings, has been a very active subject of research recently, and is now well understood using tools from various mathematical areas (see [3] for a comprehensive survey).

6 Acknowledgments

The author thanks Juanjo Rué for fruitful discussions and suggestions, and the organizers of the Workshop *Enumerative Combinatorics*, held at the *Mathematisches Forschungsinstitut Oberwolfach* from 2nd to 8th of March 2014, to which this snapshot is related.

Image credits

Fig. 6, left: [Without title (Figure 4)], Author: Richard Kenyon, in [3]

Fig. 6, right: "Aztec diamond 512×512". Author: Christopher Moore, http://tuvalu.santafe.edu/~moore/aztec512.gif [Online; accessed 05-December-2015]

References

[1] N. Elkies, G. Kuperberg, M. Larsen, and J. Propp, *Alternating-sign matrices and domino tilings (part i)*, Journal of Algebraic Combinatorics (1992), no. 1, 111–132.

[2] É. Janvresse and T. de la Rue, *Pavages aléatoires par touillage de dominoes*, 2009, http://images.math.cnrs.fr/Pavages-aleatoires-par-touillage.html, visited on December 5, 2014.

[3] R. Kenyon, *Lectures on dimers*, arXiv:0910.3129v1, 2009.

[4] C. Krattenthaler, *Another involution principle-free bijective proof of Stanley's hook-content formula*, J. Combin. Theory Ser. A **88** (1999), no. 1, 66–92.

[5] J. G. Propp and D. B. Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Random Structures Algorithms **9** (1996), no. 1-2, 223–252.

[6] ———, *How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph*, J. Algorithms **27** (1998), 170–217.

[7] J. Rué, *Domino tilings of the Aztec diamond*, Snapshots of modern mathematics from Oberwolfach $\mathcal{N}^o$16/2015, 2015.

[8] W.P. Thurston, *Conway's tiling groups*, American Mathematical Monthly **97** (1990), no. 8, 757–773.

[9] D. B. Wilson, *Mixing times of lozenge tiling and card shuffling Markov chains*, Ann. Appl. Probab. **1** (2004), no. 14, 274–325.

Éric Fusy is a CNRS researcher in the combinatorics team of LIX (computer science laboratory) at École Polytechnique, Palaiseau, France.

Mathematical subjects
Discrete Mathematics and Foundations

Connections to other fields
Computer Science, Physics

DOI
10.14760/SNAP-2016-002-EN

Snapshots of modern mathematics from Oberwolfach are written by participants in the scientific program of the Mathematisches Forschungsinstitut Oberwolfach (MFO). The snapshot project is designed to promote the understanding and appreciation of modern mathematics and mathematical research in the general public worldwide. It started as part of the project "Oberwolfach meets IMAGINARY" in 2013 with a grant by the Klaus Tschira Foundation. The project has also been supported by the Oberwolfach Foundation and the MFO. All snapshots can be found on www.imaginary.org/snapshots and on www.mfo.de/snapshots.

Mathematisches Forschungsinstitut
Oberwolfach gGmbH
Schwarzwaldstr. 9 – 11
77709 Oberwolfach
Germany

Director
Gerhard Huisken

Mathematisches Forschungsinstitut Oberwolfach
Member of the Leibniz Association
Klaus Tschira Stiftung gemeinnützige GmbH
KTS
oberwolfach FOUNDATION
IMAGINARY open mathematics