Random Walks on finite graphs and groups

Bachelorarbeit
in der Studienrichtung Mathematik

Sarah Schneeberger

Verfasst im Rahmen der Lehrveranstaltung
Seminar mit Bachelorarbeit, SE2

Betreuer: Univ.-Prof. Dr. Tim Netzer, Institut für Mathematik

Innsbruck, Oktober 2018

Contents

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.
Ich erkläre mich mit der Archivierung der vorliegenden Bachelorarbeit einverstanden.

Datum, Ort	Unterschrift

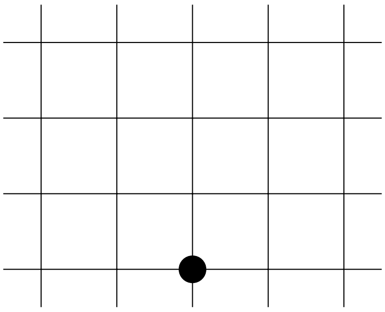
Chapter 1

Introduction

The work "Random Walks on finite Graphs and Groups" is based on the Chapter "Probability and Random Walks on Groups" of the book "Representation Theory of Finite Groups". [1]

One of the first points is to discuss what a Random Walk is. It is a process where objects move randomly - one step does not depend on the next step. With every step a decision is made whether the object will go further and if so in which direction it moves. This can be illustrated briefly by the next example.

Example 1.0.1. Imagine that there is a tourist in a city without a map. He has just left the hotel and starts wandering off with no particular destination in mind. The grid models the city where the vertices represent intersections and the edges represent streets. Each time the tourist reaches an intersection, he randomly chooses a street and continues his way.



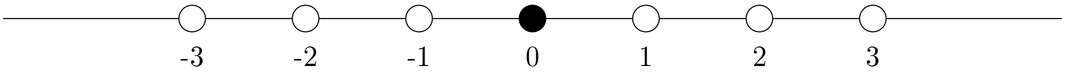
[1] Steinberg, Benjamin: Representation Theory of Finite Groups. An Introductory. Springer: New York 2012.

Questions that could be asked are:

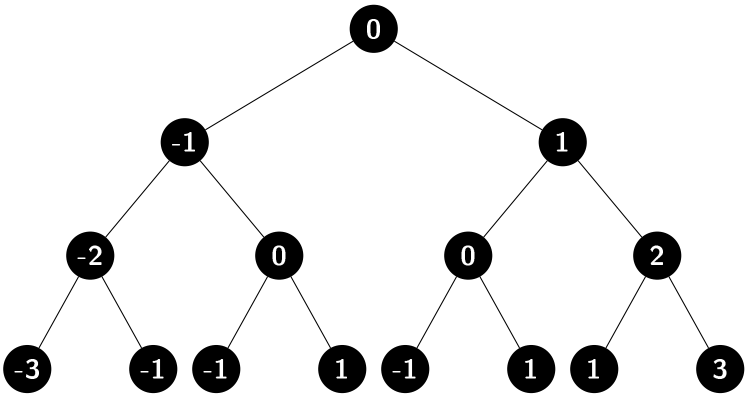
- ▷ What is the probability of returning to the starting point/hotel after n steps?
- ▷ What is the probability distribution on the intersections that describes where the tourist is after n steps?
- ▷ How long does it take before he is equally likely to be at any intersection?
- ▷ If the Graph is infinite, it is interesting to ask what is the probability of ever returning to the starting point/hotel?

This work deals with these questions in a more general form. The next example will illustrate the problem in a more mathematical way.

Example 1.0.2. There is a one-dimensional example with the infinite group $\mathbb{Z}$. At every step there is randomly chosen to go right $(+1)$ or left (-1).



Because there are always two possibilities, one has the probability $1/2$ to go left and $1/2$ to go right. This can be illustrated briefly by considering the probabilities after three steps.



One will reach position 3 or -3 with the probability of $1/8$ each and 1 or -1 with $3/8$ each.

Another example of what is meant by Random Walk is shuffling cards. It works on a model of riffle shuffling n cards as a Random Walk on the symmetric group S_n. An interesting question to ask is, how many shuffles are necessary until the cards are mixed randomly.

Of course, tourists who do not find their way and shuffling cards do not seem all that worthy of a task. But there exist many physical processes, for example a particle in a diffusion. Furthermore, there is the configurations of some objects. These represent how one configuration can be transformed after a single step and is based on the same model as the card example.

Chapter 2

Theory

The first section of this work will introduce some necessary theories.

2.1 General Theory

Definition 2.1.1 (Finite Group)
A Group $(G, *)$ is called finite group, if G has a finite amount of elements.

Definition 2.1.2 (Finite Graph and Oriented Graph [1])
A finite graph G is a sorted pair (V, E). V is a finite quantity and E is a finite quantity of two partial quantities of V. The elements of V are called vertices and the ones of E edges of G. An oriented graph is a graph in which the oriented edges (x, y) and (y, x) do not exist at the same time. The sorted pairs $(x, y) \in E$ are called oriented edges.

Concerning the graphs we often refer to the Cayley Graph therefore a definition is needed.

Definition 2.1.3 (Symmetric Generator System)
A generator system E is called symmetric generator system if $a \in E \Leftrightarrow a^{-1} \in E$.

Definition 2.1.4 (Cayley Graph [2] [3])
Let G be a group with a finite symmetric generator system $\{g_1, ..., g_n\}$, therefore $G = \langle g_1, ..., g_n \rangle$. We associate the group G with the generator system to an oriented graph $\Gamma_G(g_1, ..., g_n)$ in the following way:

[1] Cf: Matoušek, Jiří / Nešetřil, Jaroslav: Diskrete Mathematik. Eine Entdeckungsreise (2nd edition). Springer: Berlin Heidelberg 2007: pages 119 and 151.

[2] Cf: Ohshika, Ken'ichi: Translations of Mathematical Monographs. American Mathematical Society: USA 2002: page 4.

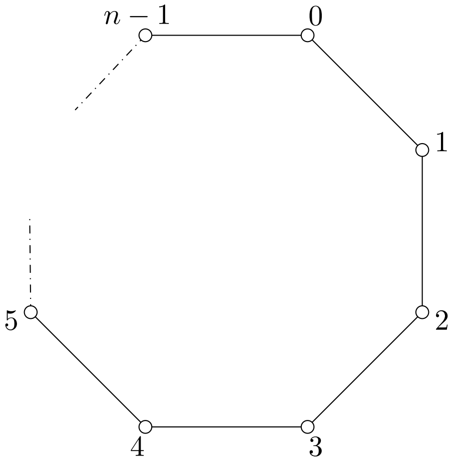
[3] Cf: Rosebrock, Stephan: Geometrische Gruppentheorie. Friedr. Vieweg&Sohn Verlag: Wiesbaden 2004: page 76.

- The points of Γ_G are the elements of G.
- The points $x, y \in G$ are joined by an edge if and only if there exists an element $g \in G$ such that $x = yg$.

Γ_G is called the Cayley Graph.

Remark 2.1.5. If Γ is the Cayley Graph of a group G, then a Random Walk on Γ is also referred to as a Random Walk on G.

Example 2.1.6. We have the finite Group $(\mathbb{Z}/n\mathbb{Z}, +)$ with the generators $\{1, -1\}$. If so, the Cayley Graph looks like this:



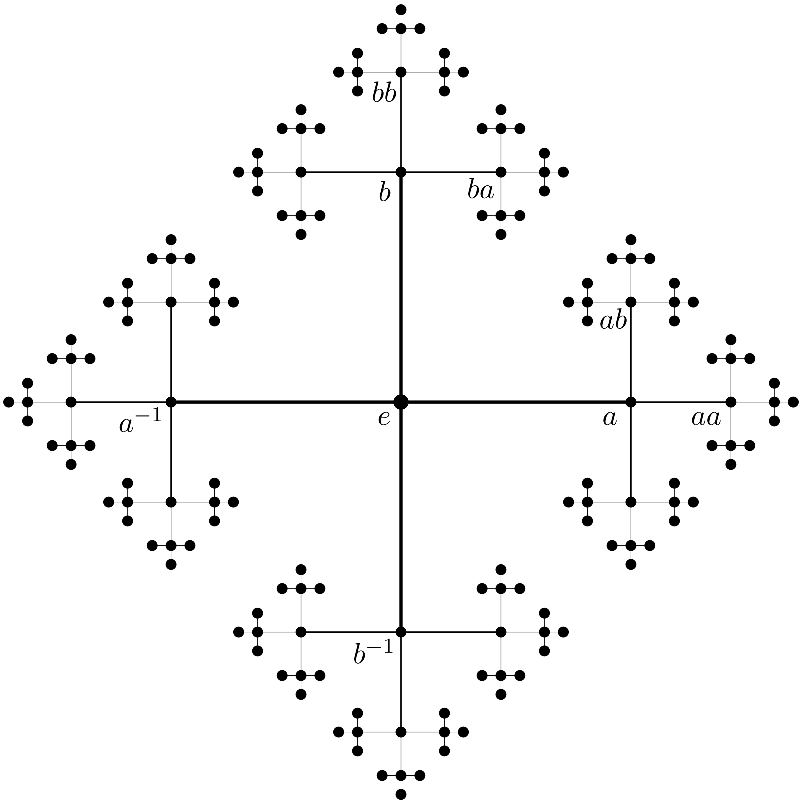
To introduce another example, a definition is needed.

Definition 2.1.7 (Free Group [4])
Let G be a group, S a quantity, and $f : S \to G$ a function. A group is called free group if it has the following universal feature: If there is a function $f' : S \to G'$ in a group G', there exists an explicit homomorphism $\varphi : G \to G'$ with $\varphi \circ f = f'$.

$$\begin{tikzcd} S \arrow[rr, "f"] \arrow[dr, "f'"'] & & G \arrow[dl, "\varphi"] \\ & G' & \end{tikzcd}$$

[4] Cf: Wüstholz, Gisbert: Algebra (2nd edition). Springer: Wiesbaden 2013: page 66.

Example 2.1.8. $(\mathbb{F}_2\left(a,b\right),*)$ is a free group with the generator system $\{a,b,a^{-1},b^{-1}\}$. The Cayley Graph of the infinite group looks like this:



We can observe that there are only the basis relations; therefore, it is not possible to reach a point with two different ways, ab is not the same as ba.

2.2 Probabilities on Groups

As already mentioned in the beginning, we consider questions about the probability. But how is the probability on groups defined? We will work with probability distributions instead of random variables.

Definition 2.2.1 (Probability Distribution)
A probability distribution on a finite group G is a function $P: G \longrightarrow [0,1]$ such that $\sum_{g\in G} P(g) = 1$ holds. If $A \subseteq G$, we put $P(A) = \sum_{g\in A} P(g)$.

Example 2.2.2. Let $G = \mathbb{Z}/2\mathbb{Z}$ be a group, so we know that through division by 2 with a remainder that the remainder is either 0 or 1, which leads to $\mathbb{Z}/2\mathbb{Z} = \{0,1\}$.

$$P([0]) = \frac{1}{2} \qquad \text{and} \qquad P([1]) = \frac{1}{2}$$

If P is a probability on G, we randomly choose an element X from G according to P, then the probability $X = g$ is given by $P(g)$. Thus, P is a probability on G for which $[0]$ and $[1]$ are equally probable. The condition from the definition is fulfilled.

Definition 2.2.3 (Support)
Let G be a finite group, P a probability on G and $P : G \longrightarrow [0,1]$ the probability distribution on G. The support of the probability P is the set $supp(P) = \{g \in G | P(g) \neq 0\}$.

Definition 2.2.4 (Uniform Distribution)
Let G be a finite group. Then the uniform distribution U on G is given by

$$U(g) = \frac{1}{|G|} \text{ for all } g \in G.$$

Definition 2.2.5 (Convolution)
Let P and Q be probabilities, then the convolution is defined as

$$(P * Q)(s) = \sum_{j \in G} P(s - j) Q(j).$$

Proposition 2.2.6. Let P and Q be probabilities on G. Then $P * Q$ is a probability on G with support $supp(P * Q) = supp(P) \cdot supp(Q)$.

Proof. $(P * Q)(g) \in [0,1]$ because

$$0 \leq \sum_{h \in G} P(g - h) Q(h) \leq \sum_{h \in G} Q(h) = 1.$$

Moreover, the following equation holds

$$\sum_{g \in G} (P * Q)(g) = \sum_{g \in G} \sum_{h \in G} P(g - h) Q(h) = \sum_{h \in G} Q(h) \sum_{g \in G} P(g - h) = \sum_{h \in G} Q(h) = 1.$$

The third equality uses that $g - h$ runs through each element of G exactly once as g does if h is fixed. It follows that $P * Q$ is a probability distribution on G. Notice that $(P*Q)(g) \neq 0$ if and only if there exists $h \in G$ such that $P(g-h) \neq 0$ and $Q(h) \neq 0$. Set $x = g-h$ and $y = h$, it follows that $(P*Q)(g) \neq 0$ if and only if there exists $x \in supp(P)$ and $y \in supp(Q)$ such that $xy = g$. Thus, $supp(P * Q) = supp(P) \cdot supp(Q)$. □

Chapter 3

Random Walks

In the introduction, the chapter term Random Walk, was visualised with consideration of common sense. Now it is necessary to clarify exactly what is meant by a Random Walk.

3.1 Calculation of a Random Walk

Definition 3.1.1 (Random Walk)
Let P be a probability on a finite group G. Then the Random Walk on G driven by P is the sequence of probability distributions $\left(P^{*k}\right)_{k=0}^{\infty}$.

Remark 3.1.2. $\left(P^{*k}\right)$ is the k-th convolution power of P, $\left(P^{*k}\right) = \underbrace{P * P * ... * P}_{\text{k}}$ and $\left(P^{*0}\right) = \delta_0$ with the Dirac-Delta-function:

$$\delta_0(P) = \begin{cases} 0 & \text{falls } 0 \notin P \\ 1 & \text{falls } 0 \in P \end{cases}$$

We can imagine it in the following way. The Random Walker starts at the identity and chooses an element X_1 from G according to the distribution and then he moves to X_1. After that, he chooses an element X_2 according to the distribution and moves to X_2X_1. We consider a sequence of independent and identically distributed random variables $X_1, X_2, ..$ with the same distribution.

The sequence of random variables forms what is called a Markov Chain in the probability literature.

Definition 3.1.3 (Markov Chain [1])
A sequence of random variables $X_0, X_1, ...$ is a Markov chain if for all $x_1, x_2, ...$

$$P\left[X_{t+1} = x_{t+1} \mid X_t = x_t \wedge X_{t-1} = x_{t-1} \wedge ... \wedge X_0 = x_0\right] = P\left[X_{t+1} = x_{t+1} \mid X_t = x_t\right]$$

is fulfilled.

The definition of a Random Walk can be illustrated shortly by the next example.

Example 3.1.4. (Discrete Circle) One considers the Cayley Graph Example 2.1.6 from the beginning. Let $0 \leq p,\, q \leq 1$ with $p+q=1$. Imagine that a particle is moving around the vertices of a regular n-gon. The particle moves one step clockwise with probability p and one step counter-clockwise with probability q. This is the Random Walk on $\mathbb{Z}/n\mathbb{Z}$ driven by the probability $p\delta_{[1]} + q\delta_{[-1]}$.

A well-known example is the Ehrenfest's urn model. It describes, for instance, an exchange of gas molecules between two cases. More generally, it can also be described with urns and balls.

Model 3.1.5. (Ehrenfest's urn) Imagine that there are two urns A and B containing a total of n balls. At the beginning, all the balls are in urn A and all of them are equally probable. At each step in time, one of the n balls is chosen at random and is moved to the other urn. At the start, all the balls are numbered consecutively from 0 to n. Let $c = (c_1, ...c_n) \in (\mathbb{Z}/2\mathbb{Z})^n$ with

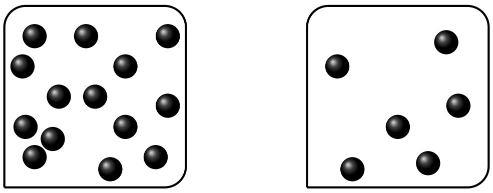
$$c_i = \begin{cases} 0 & \text{if ball } i \text{ is in urn A} \\ 1 & \text{if ball } i \text{ is in urn B} \end{cases}$$

So the initial configuration is the identity $(0, ..., 0)$. Moreover, $e_1 = (1, 0, ..., 0)\,, ..., e_n = (0, ..., 0, 1)$ is the standard basis. e_i is the vector with 1 in the i-th-coordinate and with 0 in all other coordinates for $i = \{1, ..., n\}$. Then the configuration corresponding to $e_i + c$ is obtained from the configuration corresponding to c by switching which urn contains ball i. Thus the stochastic process of switching the balls between urns corresponds to the Random Walk on $(\mathbb{Z}/2\mathbb{Z})^n$ driven by the probability

$$P = \frac{1}{n}\left(\delta_{e_1} + ... + \delta_{e_n}\right).$$

[1] Cf: Sericola, Bruno: Markov Chains. Theory, Algorithms and Applications. John Wiley & Sons: USA 2013: page 11.

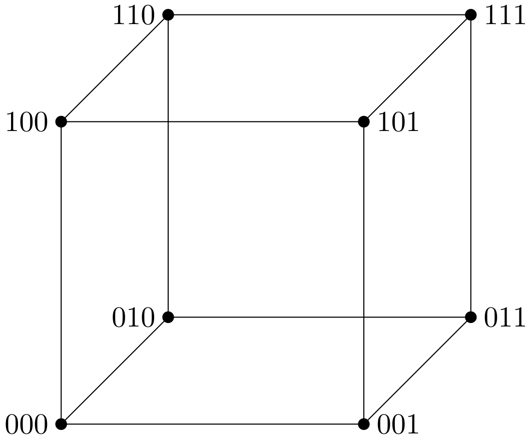
Example 3.1.6. In total there are 20 molecules. 14 are in the left box and 6 in the right one.



In the next step a molecule moves from the left to the right with the probability $^{14}/_{20} = {^7}/_{10}$ and vice versa with a probability from $^6/_{20} = {^3}/_{10}$.

Remark 3.1.7. The Ehrenfest's urn model can be considered as a simple Random Walk on the Cayley Graph of $(\mathbb{Z}/2\mathbb{Z})^n$ with respect to the symmetris set $\{e_1, ..., e_n\}$.

Example 3.1.8. In the case $n = 3$ appears a cube.
The standard basis $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ is given.
Due to $(\mathbb{Z}/2\mathbb{Z})^3$ there are 8 vertices and by adding the standard basis to the vertices one gets the edges, namely three for every vertex.



To describe the changes in every step, it also needs the transition matrix. But to define this matrix we also need the adjacency matrix.

Definition 3.1.9 (Adjacency Matrix [2])
For the adjacency matrix A of a graph $G = (V, E)$ the vertices are numbered arbitrarily

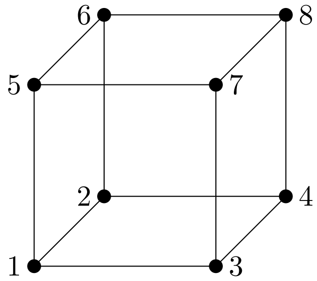
[2] Cormen, Thomas / Leiserson, Charles / Rivest, Ronald / Stein, Clifford: Algorithmen-Eine Einführung. Oldenbourg Wissenschaftsverlag GmbH: München 2013. page 601.

from 1 to $\mid V \mid$. Then the adjacency matrix is a $\mid V \mid \times \mid V \mid$ matrix $A = (a_{ij})$ with the elements

$$a_{ij} = \begin{cases} 1 & \text{if } (ij) \in E \\ 0 & \text{else} \end{cases}$$

Remark 3.1.10. The adjacency matrix describes where the edges between the vertices are. In the definition above all the edges have the same probability; otherwise, one has to take the weight of the edges into acount too.

Example 3.1.11. Let's calculate the adjacency matrix of the Ehrenfest example with three balls (Example 3.1.8). The vertices are numbered in the following way:



The adjacency matrix A looks like that one below.

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Definition 3.1.12 (Transition Matrix [3])
The transition matrix T of a Random Walk is obtained from the adjacency matrix A by $M = \frac{1}{\mid S \mid} A$ and has the following features:

- ▷ All the entries are between $[0, 1]$
- ▷ The sum of every row is 1

Remark 3.1.13. $\mid S \mid$ describes the number of generators. In other words, how many edges are possible at every vertex.

[3] Lange, Tanja / Tahagi, Tsuyoshi: Post-Quantum Cryptography. Springer: Schweiz 2017. page 72.

Example 3.1.14. The transition matrix for the Ehrenfest example (Example 3.1.8 and Example 3.1.11) is

$$T = \frac{1}{3}\begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Proposition 3.1.15. Let $v^{(0)}$ be the vector of the startdistribution and T the transition matrix. For one step the distribution is calculated by $v^{(1)} = T \cdot v^{(0)}$ and for k-steps it is $v^{(k)} = T^k \cdot v^{(0)}$.

Proof. [4] [5] Proof by induction and we use the Markov Chain (Definition 3.1.3).

k=1: We get for $j = 1, ..., n$ that

$$v_j^{(1)T} = P(X_1 = s_j) = \sum_{i=1}^{n} P(X_0 = s_i, X_1 = s_j) = \sum_{i=1}^{n} P(X_0 = s_i)P(X_1 = s_j \mid X_0 = s_i) = \sum_{i=1}^{n} v_i^{(0)T} T_{ij}$$

which is the j-th element of the row vector $v^{(0)T} \cdot T$. Hence $v^{(1)} = T \cdot v^{(0)}$.

k+1: We get for $j = 1, ..., n$ that

$$\begin{aligned} v_j^{(k+1)T} = P(X_{k+1} = s_j) = \sum_{i=1}^{n} P(X_k = s_i, X_{k+1} = s_j) &= \sum_{i=1}^{n} P(X_k = s_i)P(X_{k+1} = s_j \mid X_k = s_i) \\ &= \sum_{i=1}^{n} v_i^{(k)T} T_{ij}. \end{aligned}$$

So $v^{(k+1)} = T \cdot v^{(k)}$. But $v^{(k)} = T^k \cdot v^{(0)}$ by the induction hypothesis, so that
$v^{(k+1)} = T \cdot v^{(k)} = T^k \cdot T \cdot v^{(0)} = T^{k+1} \cdot v^{(0)}$. □

[4] Cf: Kaufmann, Edgar: Seminar über Markovketten. Universität Siegen 2003. page 5.

[5] Cf: Depperschmidt, Andrej: Markovketten. Universität Freiburg 2016. page 10.

Example 3.1.16. We consider again the Ehrenfest's urn example with three balls. Let $v^{(0)}$ be the startvector $(1, 0, 0, 0, 0, 0, 0, 0)$. Then $v^{(1)} = (0, 1/3, 1/3, 0, 1/3, 0, 0, 0)$.

$v^{(2)}$	0.3333	0	0	0.2222	0	0.2222	0.2222	0
$v^{(5)}$	0	0.251	0.251	0	0.251	0	0	0.2469
$v^{(10)}$	0.250	0	0	0.249996	0	0.249996	0.249996	0
$v^{(25)}$	0	0.25	0.25	0	0.25	0	0	0.25
$v^{(55)}$	0	0.25	0.25	0	0.25	0	0	0.25
$v^{(100)}$	0.25	0	0	0.25, 0	0.25	0.25	0	0

Table 3.1 Calculation of iterations

What stands out in the table is that the Ehrenfest's urn model with three balls does not converge to the uniform distribution. But this is the aim that one can show that every configuration is equally probable.

If we now turn to a necessary condition that every configuration can be anyway equally probable.

Definition 3.1.17 (Ergodic)
A Random Walk on a group G driven by a probability P is said to be ergodic if there exists an integer $k > 0$ such that $P^{*k}(g) > 0$ for all $g \in G$; for example, $supp(P^{*k}) = G$

Remark 3.1.18. In other words, for some k it is possible to go from any state to any state in exactly k steps.

Proposition 3.1.19. Let P be a probability on a finite group G and suppose that:

1. $P(1) > 0$
2. $supp\,(P)$ generates the group G

Then the Random Walk driven by P is ergodic.

Proof. Let $S = supp(P)$. One knows that $S^k \subseteq S^{k+1}$ for all $k \geq 0$ because $1 \in S$. Due to the fact that S is a generating set for G, there exists $N > 0$ such that $S^N = G$. By Proposition 2.2.6 the support of P^{*N} is G, that is, $P^{*N}(g) > 0$ for all $g \in G$. This is exactly the definition of ergodic. □

The next example may help to understand the term ergodic better.

Example 3.1.20. Imagine the Ehrenfest's urn model with three balls. One starts at the point 1 of the cube then after 3 steps (odd number of steps) we can reach the positions $2, 3, 8, 5$ and after 2 steps (even number of steps) we can reach the positions $1, 4, 6, 7$. The Ehrenfest's urn model is not ergodic because one cannot go from any state to any state in exactly k steps.

Remark 3.1.21. Instead of the described model one often considers the Lazy Random Walk on the Cayley Graph $(\mathbb{Z}/2\mathbb{Z})^n$.

The following model is useful because it fulfills the ergodic condition. To define the model, a definition is needed.

Definition 3.1.22 (Symmetric Subset)
Let G be a group and S its subset. It is called symmetric if $x^{-1} \in S$ whenever $x \in S$. In other words, $S = S^{-1}$.

Model 3.1.23. (Lazy Random Walk) The following model is called the Lazy Random Walk on Γ. Let S be a symmetric subset of the group G and let Γ be the corresponding Cayley Graph. At each step there are two possibilities:

- ▷ stay at the current vertex with probability 1/2
- ▷ move along an edge to a neighbour vertex with probability 1/2

. This is modelled by the Random Walk on G driven by the probability

$$P = \frac{1}{2}\delta_1 + \frac{1}{2|S|}\delta_S.$$

Remark 3.1.24. Different methods have been proposed in the literature to weight the two variant possibilities in the model above. It is just necessary that the factors sum up to one. Throughout this work, the probability $^1/_2$ is used to stay at the current vertex and $^1/_2$ to move on.

Example 3.1.25. $G = \mathbb{Z}/n\mathbb{Z}$ and $S = \{1, -1\}$ are given. We calculate the probability $|S| = 2$

$$\frac{1}{2}\delta_0 + \frac{1}{2 \cdot 2}\left(\delta_1 * \delta_{-1}\right)$$

Then the probability is

$$P = \frac{1}{2}\delta_{[0]} + \frac{1}{4}\delta_{[1]} + \frac{1}{4}\delta_{[-1]}.$$

Proposition 3.1.26. [6] Let T be a symmetric and ergodic transition matrix. Then $\lambda_1 = 1$ is a simple eigenvalue with eigenvector $(1, ..., 1)^T$ and all other eigenvalues λ_j, $j = 2, ..., n$, satisfy

$$-1 < \lambda_j < 1. \tag{3.1}$$

Proof. Let T be a symmetric and ergodic $n \times n$ matrix. $T(1, ..., 1)^T = (1, ..., 1)^T$ because the row sums are all 1 due to the fact that T is a transition matrix. Thus $\lambda_1 = 1$ is an eigenvalue with eigenvector $(1, ..., 1)^T$ and the normalized vector is $u_1 = \dfrac{(1, ..., 1)^T}{\sqrt{n}}$. What remains to be shown is that all other eigenvalues are smaller than 1. Let u_j be any nonzero eigenvector orthogonal to $(1, ..., 1)$ then $Tu_j = \lambda_j u_j$ for some eigenvalue λ_j. We know now that $(1, ..., 1)u_j = 0$ but we do not know yet that $\mid \lambda_j \mid < 1$. Because T is symmetric the startdistribution is the uniform distribution ${}^{(1,...,1)^T}\!/_n$ and so the transition matrix after k-steps T^k must converge as $k \to \infty$ to the matrix with all entries equal to ${}^1\!/_n$. Consequently, $\lim_{k\to\infty} T^k u_j = (1, ..., 1)\frac{u_j}{n} = 0$. But because $Tu_j = \lambda_j u_j$ it is $T^k u_j = \lambda_j^k u_j$. Thus, it must be the case that $\lambda_j^k \longrightarrow 0$. Therefore, the eigenvalue $\lambda_1 = 1$ is simple and all other eigenvalues are strictly less than one. □

Remark 3.1.27. Let P be a probability on a finite abelian group G and suppose that the Random Walk driven by P is ergodic. Then the sequence (P^{*k}) converges to the uniform distribution.

This remark means intuitively that an ergodic Random Walk on a finite group G can be used to randomly generate elements of G.

Example 3.1.28. Consider the Lazy Random Walk of the Ehrenfest's urn with three balls. Then the transition matrix is $\hat{T} = \frac{1}{2}I_n + \frac{1}{2}T$. In this particular case it looks like that:

$$\hat{T} = \frac{1}{2}\begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1 \end{bmatrix} + \frac{1}{2}\cdot\frac{1}{3}\begin{bmatrix} 0&1&1&0&1&0&0&0\\ 1&0&0&1&0&1&0&0\\ 1&0&0&1&0&0&1&0\\ 0&1&1&0&0&0&0&1\\ 1&0&0&0&0&1&1&0\\ 0&1&0&0&1&0&0&1\\ 0&0&1&0&1&0&0&1\\ 0&0&0&1&0&1&1&0 \end{bmatrix}$$

[6] Cf: Lalley, Steven: Convergence rates of Markov Chains.

$$= \begin{bmatrix} 1/2 & 1/6 & 1/6 & 0 & 1/6 & 0 & 0 & 0 \\ 1/6 & 1/2 & 0 & 1/6 & 0 & 1/6 & 0 & 0 \\ 1/6 & 0 & 1/2 & 1/6 & 0 & 0 & 1/6 & 0 \\ 0 & 1/6 & 1/6 & 1/2 & 0 & 0 & 0 & 1/6 \\ 1/6 & 0 & 0 & 0 & 1/2 & 1/6 & 1/6 & 0 \\ 0 & 1/6 & 0 & 0 & 1/6 & 1/2 & 0 & 1/6 \\ 0 & 0 & 1/6 & 0 & 1/6 & 0 & 1/2 & 1/6 \\ 0 & 0 & 0 & 1/6 & 0 & 1/6 & 1/6 & 1/2 \end{bmatrix}$$

From the transition matrix $\hat{T}$, it can be seen that it fulfills the conditions that all entries are between $[0,1]$ and the sum of every row is 1. The eigenvalues of the matrix are $\lambda_1 = 1$, $\lambda_{2,3,4} = {}^1\!/_3$, $\lambda_{5,6,7} = {}^2\!/_3$, $\lambda_8 = 0$. Another important finding is that there is a simple eigenvalue $\lambda_1 = 1$ and all others fulfill the inequation $-1 < \lambda_j < 1$. Therefore, the matrix converges to the uniform distribution which can also be seen in the table below.

$v^{(0)}$	1	0	0	0	0	0	0	0
$v^{(1)}$	0.5	0.16666667	0.16666667	0	0.16666667	0	0	0
$v^{(2)}$	0.33333333	0.16666667	0.16666667	0.05555556	0.16666667	0.05555556	0.05555556	0
$v^{(5)}$	0.17592593	0.1409465	0.1409465	0.10802469	0.1409465	0.10802469	0.10802469	0.07716049
$v^{(10)}$	0.13150942	0.12716557	0.12716557	0.12283019	0.12716557	0.12283019	0.12283019	0.11850328
$v^{(15)}$	0.1258564	0.12528545	0.12528545	0.12471453	0.12528545	0.12471453	0.12471453	0.12414365
$v^{(20)}$	0.12511277	0.12503759	0.12503759	0.12496241	0.12503759	0.12496241	0.12496241	0.12488723
$v^{(30)}$	0.12500196	0.12500065	0.12500065	0.12499935	0.12500065	0.12499935	0.12499935	0.12499804
$v^{(40)}$	0.12500003	0.12500001	0.12500001	0.12499999	0.12500001	0.12499999	0.12499999	0.12499997

Table 3.2 Iterations of the Lazy Random Walk

Proposition 3.1.29. [7] For a Lazy Random Walk all the eigenvalues of $\hat{T}$ are nonnegative.

Proof. $\hat{T} = \frac{1}{2}I_n + \frac{1}{2}T$ is the transition matrix of the Lazy Random Walk whereby T denotes the transition matrix of the original problem (the one without the lazyness). λ is an eigenvalue of $\hat{T}$ with the eigenvector u: $\hat{T}u = \lambda I_n u \Leftrightarrow \hat{T}u - \lambda I_n u = 0$. But this is equivalent to require that

$$(\frac{T}{2} + \frac{I_n}{2})u - \lambda I_n u = \frac{T}{2} - (\lambda - \frac{1}{2})I_n u = 0$$

holds. The inequation 3.1 implies that $2\lambda - 1 \geq -1 \Leftrightarrow \lambda \geq 0$. □

The rate of convergence of T^k as $k \to \infty$ is controlled by the nontrivial eigenvalues and their eigenvectors.

[7] Cf: Dabbs, Beau: Markov Chains and mixing times.

Proposition 3.1.30. Let T be a symmetric, ergodic $n \times n$ transition matrix with non-trivial eigenvalues $\lambda_2, \lambda_3, .., \lambda_n$ written according to multiplicity and in decreasing order of absolute value. Then the following inequation is fulfilled:

$$\| T^k v - (1/n...1/n)^T \|_2^2 \leq \lambda_2^{2k}$$

Proof. We know that $\sum_{i=1}^n v_i = 1 \Leftrightarrow \langle (1, ..., 1), v \rangle = 1$. After diagonalization T has the form

$$T = \begin{bmatrix} 1 & & & & & \\ & \lambda_2 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \lambda_n \end{bmatrix}.$$

There exists an ON-basis $((\frac{1}{\sqrt{n}} ... \frac{1}{\sqrt{n}}), v_2, ..., v_n)$ and therefore the orthogonal matrix

$$O = \begin{bmatrix} \frac{1}{\sqrt{n}} & v_2 & ... & v_n \\ \frac{1}{\sqrt{n}} & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & & & \end{bmatrix}$$

has the size n. If we transform T to O the scalar product and the norm remain.

$$\begin{gathered} e_i^T O^T T O e_j \Leftrightarrow \\ (Oe_i)^T T (Oe_j) \Leftrightarrow \\ v_i^T T v_j \Leftrightarrow \\ v_i^T \lambda_j v_j \Leftrightarrow \\ \lambda_j \langle v_i, v_j \rangle \Leftrightarrow \\ \lambda_j \delta_{ij} \end{gathered}$$

where δ_{ij} is the Kronecker-Delta. Only on the diagonal are entries.
$D = O^T T O \Rightarrow T = ODO^T$

$$\| T^k v - (1/n...1/n)^T \|_2^2$$

$$= \| OD^k O^T v - (1/n...1/n)^T \|_2^2 \; [\text{because } \langle \mathbf{1}, v \rangle = 1, \langle O^T \mathbf{1}, O^T v \rangle = \langle \sqrt{e_1}, w \rangle = \sqrt{n} w_1 \Rightarrow w = O^T v]$$

$$=\| OD^k w - O\frac{1}{\sqrt{n}}e_1 \|_2^2 = \text{[because of the fact } \| Ox \|_2 = \| x \|_2]$$

$$=\| D^k w - \frac{1}{\sqrt{n}}e_1 \|_2^2$$

$$=\| \begin{pmatrix} w_1 \\ \lambda_2^k \\ . \\ . \\ . \\ \lambda_n^k w_n \end{pmatrix} - \begin{pmatrix} 1/\sqrt{n} \\ 0 \\ . \\ . \\ . \\ 0 \end{pmatrix} \|_2^2$$

$$= \sum_{i=1}^{n} \lambda_i^{2k} w_i^2$$

$$\leq \lambda_2^{2k} \sum_{i=2}^{n} w_i^2$$

$$= \sum \lambda_2^{2k} (\| w \|_2^2 - \frac{1}{n})$$

$$\text{estimation: } \| w \|_2^2 = \| O^T v \|_2^2 = \| v \|_2^2 \leq \sum v_i = 1$$

because $v_i^2 < v_i$ due to the fact that v_i is between 0 and 1

$$\Rightarrow \lambda_2^{2k} \underbrace{(1 - \frac{1}{n})}_{\text{maximum } 1} \leq \lambda_2^{2k}$$

□

The method to calculate the eigenvalues with a matrix is particularly useful if the matrix does not have a big dimension. A major advantage of the following way to calculate the eigenvalues is that one does not need the matrix.

Definition 3.1.31 (Spectrum)
The spectrum of the Random Walk on a finite group G driven by the probability P is the set of eigenvalues of the transition matrix. It is denoted $spec(P)$.

Remark 3.1.32. In practice the generator system S is used instead of P.

Remark 3.1.33. For the Ehrenfest's urn model with $n = 3$, we get a 8×8-matrix and therefore the linear operator is $M : \mathbb{R}^8 \longrightarrow \mathbb{R}^8$.

What follows is the proposition (Diaconis) but before seeking an explanation, it is indeed needed to repeat the definitions of a group homomorphism and a character.

Definition 3.1.34 (Group Homomorphism)
Let $(G, *), (H, \circ)$ be groups. A group homomorphism is a mapping $\varphi : G \longrightarrow H$ with $\varphi(g * h) = \varphi(g) \circ \varphi(h)$ for all $g, h \in G$.

Definition 3.1.35 (Character)
A character of a finite abelian group is a homomorphism

$$\chi : G \longrightarrow (\mathbb{C} \setminus \{0\}, \cdot).$$

Proposition 3.1.36. (Diaconis) Let $G = \{g_1, ..., g_n\}$ be a finite group, $P = (p_1, ..., p_n)^T$ a probability on G, $\chi : G \longrightarrow (\mathbb{C} \setminus \{0\}, \cdot)$ a group homomorphism and $\hat{P}(\chi) := \sum \chi(g_i) \cdot p_i$. Then $spec(P) = \left\{ \hat{P}(\chi) | \chi \in \hat{G} \right\}$ holds and an orthonormal basis for the eigenspace of λ is the set of all characters χ with $\hat{P}(\chi) = \lambda$.

Proof. [8] Let T be the transition matrix and χ be a character of G.
Step 1: Then the vector $(\chi(g))_{g \in G}$ is an eigenvector of T to the eigenvalue $\hat{P}(\chi)$.

Step 2: Characters of a group are linearly independent or in other words if an equation

$$c_1\chi_1(g) + ... + c_n\chi_n(g) = 0 \tag{3.2}$$

is fulfilled in $(\mathbb{C} \setminus \{0\}, \cdot)$ for all $g \in G$, then all c_i are zero.

Proof by induction: For $n = 1$ from $c_1\chi_1(g) = 0$ results $c_1 = 0$. We can assume the thesis for $n-1$ as right. For $g, h \in G$ we get $c_1\chi_1(gh) + ... + c_n\chi_n(gh) = 0 \Leftrightarrow c_1\chi_1(g)\chi_1(h) + ... + c_n\chi_n(g)\chi_n(h) = 0$. From this equation we subtract $\chi_n(h)(c_1\chi_1(g) + ... + c_n\chi_n(g)) = 0$ and we get $c_1[\chi_1(h) - \chi_n(h)]\chi_1(g) + ... + c_{n-1}[\chi_{n-1}(h) - \chi_n(h)]\chi_{n-1}(g) = 0$. From the induction hypothesis we know that $\chi_1, ..., \chi_n$ are linearly independent this is why all coefficients must be zero: $c_i[\chi_i(h) - \chi_n(h)]\chi_i(g) = 0$ for $i = 1, ..., n-1$. Due to the fact that χ_i and χ_n are different characters we can find for every fixed i a h that $\chi_n(h) \neq \chi_i(h)$. From $c_i[\chi_i(h) - \chi_n(h)]\chi_i(g) = 0$ results $c_i = 0$ for $i = 1, ..., n-1$. We put that in equation 3.3 and get $c_n = 0$.

Step 3: For a finite abelian group with n elements, there always exists n characters.

Proof: Because of the "Struktursatz für endlich erzeugte abelsche Gruppen" we know that $G \cong \mathbb{Z} \times \mathbb{Z}/p_1^{e_1}\mathbb{Z} \times ... \times \mathbb{Z}/p_d^{e_d}$ where $p_1, ..., p_d \in \mathbb{Z}$ are prim numbers and $m, e_1, ...e_d \in \mathbb{N}$ are numbers.

Now it is necessary to proof the statement below.
Let G, H be abelian groups, then the homomorphism $\chi : G \times H \to C^x$ correspond to the homomorphisms $\chi_G : G \to C^x$ and $\chi_H : H \to C^x$.

Proof: The aim is to find a bijective mapping between $\chi : G \times H \to C^x$ and $\chi_G : G \to C^x$ and $\chi_H : H \to C^x$

[8] Cf: Van der Waerden, B.L. Algebra 1. Springer: Berlin, Heidelberg 1971. pages 159, 160.

" $\Rightarrow$ " : Let $\chi : G \times H \to C^x$ be given. Set $\chi_G : G \to C^x$ with $\chi_G(g) := \chi(g, g')$ as well as $\chi_H : H \to C^x$ with $\chi_H(h) := \chi(h', h)$ where g' and h' are the neutral elements. Both of them are then homomorphisms. Then $\chi_G(g) \cdot \chi_H(h) = \chi(g, h') \cdot \chi(g', h) = \chi((g, h') \cdot (g', h)) = \chi(g, h)$
" $\Leftarrow$ " : Let $\chi_G : G \to C^x$ and $\chi_H : H \to C^x$ be given. Set $\chi : G \times H \to C^x$ with $\chi(g, h) := \chi_G(g) \cdot \chi_H(h)$. It is a homomorphism. Then $\chi(g, h') = \chi_G(g) \cdot \chi_H(h') = \chi_G(g)$ and $\chi(g', h) = \chi_G(g') \cdot \chi_H(h) = \chi_H(h)$.
Both mappings are invers to each other and the statement is proofed.

Step 4: Because of Steps 1-3 the vectors $(\chi(g))_{g \in G}$ form a basis of eigenvectors of $\mathbb{R}^G$. □

Remark 3.1.37. From the proposition (Diaconis), we can see that the spectrum of the transition matrix T can be calculated with the group character.

By the way of illustration, the example below shows how the proposition works.

Example 3.1.38. Let's calculate the spectrum of the Example 2.1.6. Consider the Lazy Random Walk on $\mathbb{Z}/n\mathbb{Z}$ driven by the probability $P = \frac{1}{2}\delta_{[0]} + \frac{1}{4}\delta_{[1]} + \frac{1}{4}\delta_{[-1]}$. Set

$$\begin{aligned} \chi_k : \mathbb{Z}/n\mathbb{Z} &\longrightarrow (\mathbb{C}/\left\{0\right\}, \cdot) \\ 0 &\longmapsto 1 \\ 1 &\longmapsto e^{2\pi ik/n} \\ -1 &\longmapsto e^{-2\pi ik/n} \end{aligned}$$

Then the eigenvalues $\hat{P}\left(\chi_k\right)$ are calculated by

$$\hat{P}\left(\chi_k\right) = \frac{1}{2} + \frac{1}{4}\left[e^{\frac{-2\pi ik}{n}} + e^{\frac{2\pi ik}{n}}\right] =$$
$$\frac{1}{2} + \frac{1}{4}\left[\cos\left(\frac{2\pi k}{n}\right) - i\sin\left(\frac{2\pi k}{n}\right) + cos\left(\frac{2\pi k}{n}\right) + i\sin\left(\frac{2\pi k}{n}\right)\right] =$$
$$\frac{1}{2} + \frac{1}{4}2\cos\left(\frac{2\pi k}{n}\right) = \frac{1}{2} + \frac{1}{2}\cos\left(\frac{2\pi k}{n}\right).$$

More specific, for $(\mathbb{Z}/8\mathbb{Z}, +)$ is $n = 8$.

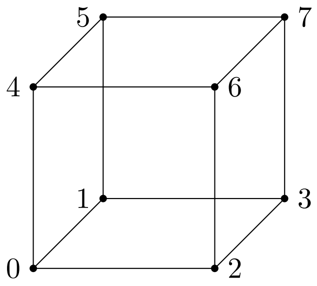
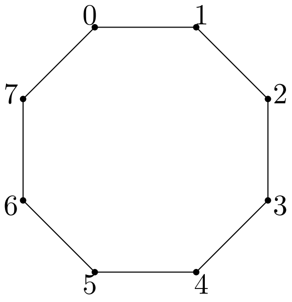
$$\begin{aligned} k = 0 &\Longrightarrow 1 \\ k = 1 &\Longrightarrow \frac{2+\sqrt{2}}{4} \\ k = 2 &\Longrightarrow \frac{1}{2} \end{aligned}$$

$$\begin{aligned} k = 3 &\Longrightarrow \frac{2-\sqrt{2}}{4} \\ k = 4 &\Longrightarrow 0 \\ k = 5 &\Longrightarrow \frac{2-\sqrt{2}}{4} \\ k = 6 &\Longrightarrow \frac{1}{2} \\ k = 7 &\Longrightarrow \frac{2+\sqrt{2}}{4} \end{aligned}$$

From the calculation above we can see that the eigenvalues and the eigenvalues of the octagon example in the next section are the same.

3.2 Comparison of a cube and an octagon

In order to compare the speed of the convergence to the uniform distribution Proposition 3.1.30 was defined. This can be seen with the two groups below.

$(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}, +)$	$(\mathbb{Z}/8\mathbb{Z}, +)$
generators: $\{(1,0),(0,1),(0,3)\}$	generators: $\{-1, 1\}$
	

The cube differs from the octagon because there are three edges from the vertices instead of just two. We can assume that the cube converge faster as the octagon due to the pictures. The aim is now to confirm this assumption.

Octagon:
The transition matrix is calculated in the same way as before explained and looks like that:

$$T = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

The eigenvalues are $\lambda_1 = 1$, $\lambda_{2,3} = 0$, $\lambda_{4,5} = \sqrt{2}/2$, $\lambda_{6,7} = -\sqrt{2}/2$, $\lambda_8 = -1$ and they do not fulfill the conditions of the Propostion 3.1.26.

$v^{(0)}$	1	0	0	0	0	0	0	0
$v^{(1)}$	0	0.5	0	0	0	0	0	0.5
$v^{(10)}$	0.265625	0	0.25	0	0.234375	0	0.25	0
$v^{(55)}$	0	0.25	0	0.25	0	0.25	0	0.25
$v^{(100)}$	0.25	0	0.25	0	0.25	0	0.25	0
$v^{(555)}$	0	0.25	0	0.25	0	0.25	0	0.25

Table 3.3 Random Walk of the octagon

This table confirms that it does not converge to the uniform distribution and we need again the Lazy Random Walk.

The transition matrix of the Lazy Random Walk is

$$\hat{T} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} + \frac{1}{2} \cdot \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1/2 & 1/4 & 0 & 0 & 0 & 0 & 0 & 1/4 \\ 1/4 & 1/2 & 1/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/4 & 1/2 & 1/4 \\ 1/4 & 0 & 0 & 0 & 0 & 0 & 1/4 & 1/2 \end{bmatrix}$$

The eigenvalues are: $\lambda_1 = 1$, $\lambda_{2,3} = 1/2$, $\lambda_{4,5} = \dfrac{-\sqrt{2}+2}{4}$, $\lambda_{6,7} = \dfrac{\sqrt{2}+2}{4}$, $\lambda_8 = 0$ and therefore it converges to the uniform distribution.

Cube:

From the Ehrenfest's urn model with three balls we know that the second largest eigenvalue is $2/3$ for the cube (Example 3.1.28).

Comparison:

In comparison the second largest of the octagon is $\dfrac{\sqrt{2}+2}{4}$. Due to Proposition 3.1.30 we can say that the cube really converges faster than the octagon.

$v^{(0)}$	1	0	0	0	0	0	0	0
$v^{(1)}$	0.5	0.25	0	0	0	0	0	0.25
$v^{(2)}$	0.375	0.25	0.0625	0	0	0	0.0625	0.25
$v^{(10)}$	0.17655945	0.1612854	0.12475586	0.0887146	0.07392883	0.0887146	0.12475586	0.1612854
$v^{(15)}$	0.14825642	0.14143938	0.12499237	0.10856062	0.10175884	0.10856062	0.12499237	0.14143938
$v^{(20)}$	0.13553328	0.13244799	0.12499976	0.11755201	0.1144672	0.11755201	0.12499976	0.13244799
$v^{(30)}$	0.12716203	0.12652878	0.125	0.12347122	0.12283797	0.12347122	0.125	0.12652878
$v^{(40)}$	0.12544378	0.1253138	0.125	0.1246862	0.12455622	0.1246862	0.125	0.1253138

Table 3.4 Calculation of the octagon

$v^{(0)}$	1	0	0	0	0	0	0	0
$v^{(1)}$	0.5	0.16666667	0.16666667	0	0.16666667	0	0	0
$v^{(2)}$	0.33333333	0.16666667	0.16666667	0.05555556	0.16666667	0.05555556	0.05555556	0
$v^{(5)}$	0.17592593	0.1409465	0.1409465	0.10802469	0.1409465	0.10802469	0.10802469	0.07716049
$v^{(10)}$	0.13150942	0.12716557	0.12716557	0.12283019	0.12716557	0.12283019	0.12283019	0.11850328
$v^{(15)}$	0.1258564	0.12528545	0.12528545	0.12471453	0.12528545	0.12471453	0.12471453	0.12414365
$v^{(20)}$	0.12511277	0.12503759	0.12503759	0.12496241	0.12503759	0.12496241	0.12496241	0.12488723
$v^{(30)}$	0.12500196	0.12500065	0.12500065	0.12499935	0.12500065	0.12499935	0.12499935	0.12499804
$v^{(40)}$	0.12500003	0.12500001	0.12500001	0.12499999	0.12500001	0.12499999	0.12499999	0.12499997

Table 3.5 Calculation of the cube

Once again, it can be seen that the cube converges faster than the octagon. The cube needs approximately 15 iterations and in contrast the octagon needs 40 iterations.

3.3 Card Shuffling

As mentioned in the introduction chapter, card shuffling describes the configurations of some objects. Moreover, it is viewed as a Random Walk on the symmetric group.

Definition 3.3.1 (Symmetric Group)
A bijective function from $\{1, 2, ..., n\}$ to $\{1, 2, ..., n\}$ is called permutation of the numbers $1, 2, ..., n$. The group of all permutations is called symmetric group S_n.

3.3.1 Top-to-Random [9]

Imagine you have a deck of n cards. Shuffling the cards by repeatedly removing the top card and returning it to a random position in the deck. All the n positions are equally probable. It is a Random Walk on the group S_n of $n!$ possible permutations of the cards. Placing the top card into the top position corresponds to the identity map. For $i \geq 2$, putting the top card in the i-th position shifts the previous positions up one and so corresponds to the cycle $(i\ i-1...1)$. Each permutation of the remaining cards is equally likely. Thus, the top-to-random shuffle has the probability

$$P = \frac{1}{n}\delta_{Id} + \sum_{i=2}^{n} \frac{1}{n}\delta_{(i\ i-1...1)}$$

Remark 3.3.2. The permutations $(2\ 1)$ and $(n\ n-1...1)$ generate S_n, this is why the walk is ergodic by Proposition 3.1.19. Due to that we can say, that after enough top-to-random shuffles the deck will be mixed.

3.3.2 Random Transpositions [10]

Consider a deck of n cards numbered from 1 to n. A configuration is a permutation of the cards in a row on a table. There are $n!$ configurations and the initial one correspond to the placement $1, 2, ..., n$. The dealer randomly picks a card with each of his hands. It is possible that the left and the right hand pick the same card. He then swaps the two cards and so if he picked the same card with each hand, then he does nothing. Such an event may occur with probability $1/n$. If he has chosen two different cards, he switches them. Given two positions $i \neq j$, there are two ways the dealer can pick this pair either the left hand picks i and the right hand picks j or vice versa. In other words, denoting by $T = \{(i, j) : i, j \in \{1, 2, ..., n\}, i \neq j\}$ the set of all unordered pairs of cards,

[9] Cf: Levin, David / Peres, Yuval: Markov Chains and Mixing Times (2nd edition). American Mathematical Society: USA 2017. pages 75 and 76.

[10] Cf: Ceccherini-Silberstein, Tullio / Scarabotti, Fabio / Tolli,Filippo: Harmonic Analysis on Finite Groups. Representation Theory, Gelfand Pairs and Markov Chains. Cambridge University Press: New York 2008. page 8.

with probability $^1/_n$ we leave the configuration unchanged, while with probability $^{(n-1)}/_n$ we randomly pick one of the $^{n(n-1)}/_2$ elements in T. At each step the probability of performing the transposition $(i\ j)$ is $^2/_{n^2}$. The probability that the dealer picks position i with both hands is $^1/_{n^2}$. Thus, the random transpositions shuffle is the random walk on S_n driven by the probability Q defined by

$$Q(\sigma) = \begin{cases} 1/n & \sigma = Id \\ 2/n^2 & \sigma \text{ is a transposition} \\ 0 & \text{else} \end{cases}$$

Remark 3.3.3. The transpositions generate S_n due to Proposition 3.1.19 this is an ergodic Random Walk and so again this shuffle will randomize the deck.

3.3.3 Riffle Shuffle [11] [12] [13] [14] [15]

Nobody really shuffles a deck by randomly swapping cards, but instead the most commonly used shuffle in practice is the riffle shuffle. In this shuffle, the dealer takes the deck, splits it into two parts and then places the top half of the pack in his right hand and the bottom half in his left hand. He then drops cards from each packet, interleaving the two packets. In a perfect shuffle, the dealer would drop one card from each packet in alternation, but in reality several cards from the same packet are often dropped at a time.

Remark 3.3.4. The model of riffle shuffling was proposed by Gilbert and Shannon and independently by Reeds. It is known as the Gilbert-Shannon-Reeds shuffle.

Definition 3.3.5 (Rising sequence)
A rising sequence in a permutation σ of $\{1, ..., n\}$ is a maximal increasing, consecutive subsequence in the image sequence $\sigma(1), \sigma(2), ..., \sigma(n)$.

Example 3.3.6. $(1, 6, 2, 3, 7, 8, 4, 5)$ has two rising sequences:

- ▷ 1,2,3,4,5
- ▷ 6,7,8

[15] Cf: Lee, Jieun / Lewandoski, Janine / Refior, Kevin: Exploring the Geometric Model of Riffle Shuffling. August 2014. pages 5,6,7,14.

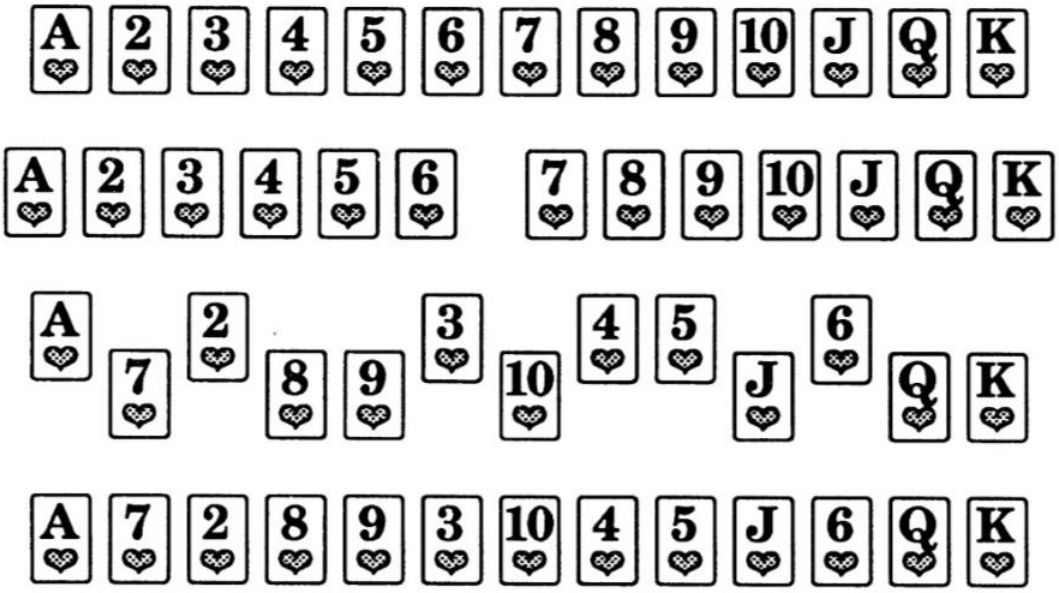


Figure 3.1 Riffle Shuffle

A riffle shuffle performs a certain permutation on the cards in the deck, which we initially assume to be labelled from 1 *to* n. The deck has n cards where 1 is the top card. The riffle shuffles correspond exactly to the permutations $\sigma \in S_n$ such that the sequence $\sigma(1), \sigma(2), ..., \sigma(n)$ consists of two rising sequences: $\sigma(1), ..., \sigma(k)$ and $\sigma(k+1), ..., \sigma(n)$.

Example 3.3.7. Imagine a deck has nine cards ordered from the top to the bottom in the following way: $2, 3, 4, 5, 6, 7, 8, 9, 10$. We perform a riffle shuffle with four cards $2, 3, 4, 5$ in the top packet and five cards $6, 7, 8, 9, 10$ in the bottom packet. Let T stands for top and B for bottom and the sequence of drops is T,B,B,T,B,T,T,B,B. In the shuffled deck the order of cards will be again, from top to bottom $6, 7, 2, 3, 8, 4, 9, 10, 5$ which corresponds to the permutation σ of the positions with image sequence $3, 4, 6, 9, 1, 2, 5, 7, 8$ as the card in the first position, 2, is now in the third position. The rising sequences of σ are $3, 4, 6, 9$ and $1, 2, 5, 7, 8$.

Remark 3.3.8. If P is the probability distribution on S_n corresponding to a Gilbert-Shannon-Reeds shuffle, then $P(\sigma) = 0$ unless σ has at most two rising sequences.

If the pack is split such that the top k cards are taken into the right hand $(0 \leq k \leq n)$ and the other $n-k$ cards into the left hand. The position of the cut is assumed to be a binomial random variable because there are $\binom{n}{k}$ ways to interleave the two hands. So if

X is the random variable counting the number of cards in the top half of the deck after the cut, then the equation holds

$$P\left[X = k\right] = \binom{n}{k}\frac{1}{2^n}$$

Notice that the closer k is to ${}^{n}\!/_{2}$ the more likely $X = k$ is to happen.

In the next step the dealer drops cards from the left and right hand until both hands are empty with the probability of dropping a card from a given hand being proportional to the number of cards in that hand. There are a cards in the right hand and b cards in the left hand. Then the probability of dropping a card from the right hand is ${}^{a}\!/_{(a+b)}$ and the probability of dropping a card from the left hand is ${}^{b}\!/_{(a+b)}$.

Proposition 3.3.9. There are k cards in the top half of the deck after the cut and n positions to place these k cards. Once these positions are chosen, the ordering of the cards is known. Thus, there are $\binom{n}{k}$ possible permutations obtainable given k. Each of them is then equally probable.

Proof. The probability of T is ${}^{a}\!/_{(a+b)}$ and of B is ${}^{b}\!/_{(a+b)}$ where a is the number of cards in the top packet and b is the number of cards in the bottom packet. So when the cut is k cards in the top packet and $n-k$ cards in the bottom packet then the probability that the sequence of drops starts with T is ${}^{k}\!/_{n}$ and the probability that the sequence starts with a B is ${}^{(n-k)}\!/_{n}$. Regardless of wheter the first drop is a T or a B, the denominator of the probability for the second drop will be reduced by 1 to $n-1$ and the numerator for the one chosen in the first position will be reduced by 1 and be either $k-1$ or $n-k-1$ according to whether the top or bottom card was dropped respectively. This will continue for each succeeding position so the numbers $k, k-1, ..., 1$ and $n-k, n-k-1, ..., 1$ will all appear in some one of the numerators. Since the probability of any sequence of drops is the product of these individual probabilities, it will be in every case ${}^{k!(n-k)!}\!/_{n!}$ and so they are all equally probable. □

Example 3.3.10. Imagine we have six cards and $k = 3$. In this case holds

$$\frac{k!\,(n-k)!}{n!} = \frac{3!\cdot 3!}{6!} = \frac{1}{20}.$$

First, we consider the sequence T,B,B,T,T,B. As mentioned before, k is the number of cards in the top half of the deck. According to the model its probability is

$$\frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{20}$$

In the beginning, we have ${}^{3}\!/_{6}$ because we start with top and there are three cards. The next element of the sequence is bottom and there are also three cards but the total amount is now five cards because we do not need to consider the first one anymore. Then there is again top and one card less on the bottom side and so on. All sequences have the probability $1/20$.

Definition 3.3.11 (Gilbert-Shannon-Reeds Shuffle)
The Gilbert-Shannon-Reeds shuffle corresponds to the Random Walk on S_n driven by the probability distribution P given by

$$P(\sigma) = \begin{cases} (n+1)/2^n & \sigma = Id \\ 1/2^n & \sigma \text{ has exactly two rising sequences} \\ 0 & \text{else} \end{cases}$$

Example 3.3.12. Let's consider a riffle shuffle over S_3. The transition matrix is a $n! \times n!$-matrix which is formed from the probabilities of transitioning from one deck to another. For every deck there are 2^n shuffles and for this example the shuffles are:

$$000, 001, 010, 011, 100, 101, 110, 111.$$

Each shuffle has a corresponding permutation and there are $n!$ permutations for every deck. In this case $123, 132, 213, 231, 321, 312$.

For shuffle (001) the corresponding permutation can be calculated as following by constructing a table with 4 columns:
Step 1: Write the identity permutation $(1, 2, ..., n)$ in column 1.
Step 2: Write down the shuffle starting with the $0's$ on the top and then the $1's$.
Step 3: Write down the shuffle with the $1's$ on the top and then the $0's$.
Step 4: Match the topmost 0 from column 2 to the topmost 0 from column 3. Continue this to the second topmost 0 and then follow by 1.
Step 5: We can see how the $0's$ and $1's$ moved to other rows and column 4 is then the corresponding permutation.

1	0 $\searrow$	1	3
2	0 $\searrow$	0	1
3	1 $\nearrow$	0	2

Table 3.6 Calculation of the corresponding permutation

We can calculate this for every shuffle and then we get the tableau:

Shuffle	Permutation	Probability
000	123	1/8
001	312	1/8
010	132	1/8
011	231	1/8
100	123	1/8
101	213	1/8
110	123	1/8
111	123	1/8

$\Rightarrow$

Permutation	Probability
123	1/2
132	1/8
213	1/8
231	1/8
312	1/8
321	0

The transition matrix is

$$T = \left[\begin{array}{ccccccc} t_{ij} & (123) & (213) & (231) & (132) & (312) & (321) \\ (123) & 1/2 & 1/8 & 1/8 & 1/8 & 1/8 & 0 \\ (213) & 1/8 & 1/2 & 1/8 & 1/8 & 0 & 1/8 \\ (231) & 1/8 & 1/8 & 1/2 & 0 & 1/8 & 1/8 \\ (132) & 1/8 & 1/8 & 0 & 1/2 & 1/8 & 1/8 \\ (312) & 1/8 & 0 & 1/8 & 1/8 & 1/2 & 1/8 \\ (321) & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/2 \end{array}\right]$$

The eigenvalues of the transition matrix are $\lambda_1 = 1$, $\lambda_{2,3,4} = {}^1\!/_2$, $\lambda_{5,6} = {}^1\!/_4$. This example fulfills the conditions of Proposition 3.1.26.
Let $(1, 0, 0, 0, 0, 0)$ be the startdistribution. Then we can calculate $v^{(k)} = T^k \cdot v^{(0)}$:

$v^{(0)}$	1	0	0	0	0	0
$v^{(1)}$	0.5	0.125	0.125	0.125	0.125	0
$v^{(2)}$	0.3125	0.15625	0.15625	0.15625	0.15625	0.0625
$v^{(3)}$	0.234375	0.1640625	0.1640625	0.1640625	0.1640625	0.109375
$v^{(4)}$	0.19921875	0.16601562	0.16601562	0.16601562	0.16601562	0.13671875
$v^{(5)}$	0.18261719	0.16650391	0.16650391	0.16650391	0.16650391	0.15136719
$v^{(6)}$	0.17456055	0.16662598	0.16662598	0.16662598	0.16662598	0.15893555
$v^{(7)}$	0.17059326	0.16665649	0.16665649	0.16665649	0.16665649	0.16278076
$v^{(8)}$	0.16862488	0.16666412	0.16666412	0.16666412	0.16666412	0.16471863
$v^{(9)}$	0.1676445	0.16666603	0.16666603	0.16666603	0.16666603	0.16569138
$v^{(10)}$	0.16715527	0.16666651	0.16666651	0.16666651	0.16666651	0.1661787
$v^{(11)}$	0.16691089	0.16666663	0.16666663	0.16666663	0.16666663	0.16642261
$v^{(12)}$	0.16678876	0.16666666	0.16666666	0.16666666	0.16666666	0.16654462

Table 3.7 Iterations of a Riffle Shuffle

With respect to the calculation above, it was found that one need about eight riffle-shuffles to mix the cards randomly.

Chapter 4

Conclusion

To sum up we can now answer the questions from the introduction chapter.

- ▷ What is the probability of returning to the starting point/hotel after k steps?
- ▷ What is the probability distribution on the intersections that describes where the tourist is after k steps?
- ▷ How long does it take before he is equally likely to be at any intersection?

The arguments given above prove that Proposition 3.1.15 including the argument $v^{(k)} = T^k \cdot v^{(0)}$ answers the first two questions. As mentioned during the work, there are necessary condition that every configuration can be equally probable. Furthermore, it is possible to calculate how fast it converges to the uniform distribution with the Proposition 3.1.30 including the argument $\| T^k v - (1/n ... 1/n)^T \|_2^2 \leq \lambda_2^{2k}$.

Bibliography

[1] Steinberg, Benjamin: Representation Theory of Finite Groups. An Introductory. Springer: New York 2012.

[2] Matoušek, Jiří / Nešetřil, Jaroslav: Diskrete Mathematik. Eine Entdeckungsreise (2nd edition). Springer: Berlin Heidelberg 2007: pages 119 and 151.

[3] Ohshika, Ken'ichi: Translations of Mathematical Monographs. American Mathematical Society: USA 2002: page 4. Online ressource (23 march 2018): `https://books.google.at/books?id=ufQEgK9pFF0C&pg=PA4&dq=cayley+graph+definition&hl=de&sa=X&ved=0ahUKEwi7-PCJ2ofaAhXII1AKHS4ND8kQ6AEIXTAH#v=onepage&q=cayley%20graph%20definition&f=false`

[4] Rosebrock, Stephan: Geometrische Gruppentheorie. Friedr. Vieweg&Sohn Verlag: Wiesbaden 2004: page 76. Online ressource (25 march 2018): `https://books.google.at/books?id=spfzBQAAQBAJ&printsec=frontcover&dq=geometrische+gruppentheorie+ein+einstieg+mit+dem+computer+basiswissen+f%C3%BCr&hl=de&sa=X&ved=0ahUKEwjbw4-ljJbaAhUOZVAKHZyjDTgQ6AEIJzAA#v=onepage&q=geometrische%20gruppentheorie%20ein%20einstieg%20mit%20dem%20computer%20basiswissen%20f%C3%BCr&f=false`

[5] Wüstholz, Gisbert: Algebra (2nd edition). Springer: Wiesbaden 2013: page 66. Online ressource (14 april 2018): `https://books.google.at/books?id=uYUkBAAAQBAJ&pg=PA67&lpg=PA67&dq=freie+gruppe&source=bl&ots=XYqTfkXdDP&sig=ElM9sPM0klhZX8FqiS_Ma85JqSw&hl=de&sa=X&ved=0ahUKEwjzlqisrLraAhVsJJoKHa_BASIQ6AEIfzAL#v=onepage&q=freie%20gruppe&f=false`

[6] Sericola, Bruno: Markov Chains. Theory, Algorithms and Applications. John Wiley & Sons: USA 2013: page 11. Online ressource (3 september 2018): `https://books.google.at/books?id=tRdwAAAAQBAJ&`

printsec=frontcover&dq=markov+chain&hl=de&sa=X&ved=0ahUKEwiK4survp7dAhXFL1AKHaMvAvcQ6AEIQjAD#v=onepage&q=markov%20chain&f=false

[7] Cormen, Thomas / Leiserson, Charles / Rivest, Ronald / Stein, Clifford: Algorithmen-Eine Einführung. Oldenbourg Wissenschaftsverlag GmbH: München 2013. page 601. Online ressource (30 may 2018):
https://books.google.at/books?id=O1pSDgAAQBAJ&pg=PA601&dq=adjazenzmatrix&hl=de&sa=X&ved=0ahUKEwi-vryj8pnbAhVIFywKHcJqAYOQ6AEIVzAI#v=onepage&q=adjazenzmatrix&f=false

[8] Lange, Tanja / Tahagi, Tsuyoshi: Post-Quantum Cryptography. Springer: Schweiz 2017. page 72. Online ressource (21 may 2018):
https://books.google.at/books?id=Yv4nDwAAQBAJ&printsec=frontcover&dq=post+quantum+cryptography&hl=de&sa=X&ved=0ahUKEwiI7cvDk5rbAhUL2SwKHcsjCc4Q6AEILTAB#v=onepage&q=post%20quantum%20cryptography&f=false

[9] Kaufmann, Edgar: Seminar über Markovketten. Universität Siegen 2003. page 5. Online ressource (1 june 2018):
http://www.uwenowak.de/arbeiten/markov.pdf

[10] Depperschmidt, Andrej: Markovketten. Universität Freiburg 2016. page 10. Online ressource (1 june 2018):
https://www.stochastik.uni-freiburg.de/mitarbeiter/depperschmidt/inhalte/markovketten.pdf

[11] Lalley, Steven: Convergence rates of Markov Chains. Online ressource (6 june 2018):
https://galton.uchicago.edu/~lalley/Courses/313/ConvergenceRates.pdf

[12] Dabbs, Beau: Markov Chains and mixing times. Online ressource (6 june 2018):
http://www.math.uchicago.edu/~may/VIGRE/VIGRE2009/REUPapers/Dabbs.pdf

[13] Van der Waerden, B.L. Algebra 1. Springer: Berlin, Heidelberg 1971. pages 159, 160. Online ressource (22 may 2018):
https://books.google.at/books?id=E4t_BwAAQBAJ&pg=PP5&dq=Algebra+I+von+Bartel+Eckmann+L.+Van+der+van+der+Waerden&hl=de&sa=X&ved=0ahUKEwjzyYGHq7zbAhWSLlAKHZgXBfgQ6AEINTAC#v=onepage&q=Algebra%20I%20von%20Bartel%20Eckmann%20L.%20Van%20der%20van%20der%20Waerden&f=false

[14] Levin, David / Peres, Yuval: Markov Chains and Mixing Times (2nd edition). American Mathematical Society: USA 2017. pages 75 and 76. Online ressource (21 may 2018):
https://books.google.at/books?id=f208DwAAQBAJ&pg=PA75&dq=top+to+random+shuffle&hl=de&sa=X&ved=0ahUKEwitx4eA-9faAhVQUlAKHUvZAB8Q6AEIYDAH#v=onepage&q=top%20to%20random%20shuffle&f=false

[15] Ceccherini-Silberstein, Tullio / Scarabotti, Fabio / Tolli, Filippo: Harmonic Analysis on Finite Groups. Representation Theory, Gelfand Pairs and Markov Chains. Cambridge University Press: New York 2008. page 8. Online ressource (21 may 2018):
https://books.google.at/books?id=3trud4weQ3AC&pg=PA10&dq=random+transpositions&hl=de&sa=X&ved=0ahUKEwjXvIWlqNjaAhUNPFAKHfWmDZgQ6AEIXTAH#v=onepage&q=random%20transpositions&f=false

[16] Hesse, Christian: Angewandte Wahrscheinlichkeitstheorie. Friedr. Vieweg&Sohn Verlag: Braunschweig/Wiesbaden 2003. page 129. Online ressource (21 may 2018):
https://books.google.at/books?id=UEj0BQAAQBAJ&pg=PA129&dq=riffle+shuffle&hl=de&sa=X&ved=0ahUKEwi5y0up69naAhUDYlAKHZF1AR8Q6AEIQDAD#v=onepage&q=riffle%20shuffle&f=false

[17] Aigner, Martin / Ziegler, Günter: Proofs from the book (3rd edition): Springer: Berlin/Heidelberg 2004. pages 163 and 164. Online ressource (21 may 2018):
https://books.google.at/books?id=KvQr9l0wgf8C&pg=PA163&dq=riffle+shuffle&hl=de&sa=X&ved=0ahUKEwi5y0up69naAhUDYlAKHZF1AR8Q6AEIUDAF#v=onepage&q=riffle%20shuffle&f=false

[18] Bayer, Dave/ Diaconis, Persi: Trailing the dovetail shuffle to its lair (Vol. 2): Columbia University and Harvard University 1992. pages 295, 296, 297, 298. Online ressource (21 may 2018):
https://books.google.at/books?id=Vt-_HgAACAAJ&dq=Trailing+the+dovetail+shuffle+to+its+lair&hl=de&sa=X&ved=0ahUKEwjCiqaq-ZzdAhXFbVAKHaRyBAMQ6AEIKjAA

[19] Hammarström, Harald: Card-Shuffling Analysis with Markov Chains. January 2015. pages 3,4,8. Online ressource (7 june 2018):
https://pdfs.semanticscholar.org/24dc/3948bd857a113ba4fa80cd241f6f60c3133a.pdf

[20] Lee, Jieun / Lewandoski, Janine / Refior, Kevin: Exploring the Geometric Model of Riffle Shuffling. August 2014. pages 5,6,7,14. Online ressource (7 june 2018): https://pdfs.semanticscholar.org/8fc9/9368932c946585ffaebea4cc86d78ec5f7fb.pdf

REGISTER OF ILLUSTRATION:

[21] **Figure 3.1 Riffle Shuffle:**
Bayer, Dave/ Diaconis, Persi: Trailing the dovetail shuffle to its lair (Vol. 2): Columbia University and Harvard University 1992. page 295. Online ressource (21 may 2018):
https://books.google.at/books?id=Vt-_HgAACAAJ&dq=Trailing+the+dovetail+shuffle+to+its+lair&hl=de&sa=X&ved=0ahUKEwjCiqaq-ZzdAhXFbVAKHaRyBAMQ6AEIKjAA