RANDOM WALKS ON GROUPS: STRONG UNIFORM TIME APPROACH

Igor Pak

Harvard University
E-mail: pak@math.harvard.edu

May 14, 1997

Abstract. In this work we develop further and apply the strong uniform time approach for study of the random walks on finite groups. This approach was introduced by Aldous and Diaconis in [AD1,2] and later developed by Diaconis and Fill in [DF].

Consider a random walk on a finite group. We introduce a notion of *total separation* for this walk which can be thought as a new measure of how fast the walk is mixing. We show that the total separation is the mean of the best possible strong uniform time. We prove various bounds on the total separation, find connections with hitting times and establish relations between total separations under several natural operations on walks on groups, such as rescaling of the walk, taking direct and wreath product of groups.

In this work the emphasis is given to the study of concrete examples of walks. The successful applications of the method include finding sharp bounds on the total separation for the natural random walks on cube, cyclic group, dihedral group, symmetric group, hyperoctahedral group, Heisenberg group, and others. In several cases we were able to obtain not only sharp bounds, but find the exact value of the total separation.

Key words and phrases. Random walk, Markov chain, stopping time, strong uniform time, hitting time, separation distance, symmetric group, Heisenberg group, full linear group, finite field, Cayley graph, diameter, randomized algorithm.

Introduction

In this thesis we use a strong uniform time approach to study random walks on finite groups. We present explicit constructions of such times for the random walks on a cycle, cube, for various walks on symmetric groups, for upper triangular matrices, and for other nice special cases.

Before I explain what the strong uniform time approach is let me briefly recall some historical bright-spots related to the thesis. The story began in the pioneer works by Markov who had in mind important examples of card shuffles when introducing Markov chains. The next major step was made by Polya who studied the random walk on $\mathbb{Z}$. Although not finite, this walk is related to the random walk on a cycle which was noted by subsequent investigators.

In the eighties the theory of random walks on finite groups obtained its independence, its own problems and techniques. In the pioneer paper [DSh1] Diaconis and Shahshahani introduced the technique which involves bounding of the irreducible characters of the group. They were able to apply this technique to the problem of shuffling a deck of cards by switching a randomly chosen pair of cards. It turns out that this random walk can be thought as a walk on a symmetric group generated by a single conjugacy class, so one can use the existing theory of characters of S_n to obtain the estimates on the convergence. Later on this technique was applied to get sharp bounds for the walks on a wide variety of groups (see e.g. [DSh2], [Lulov], [Gluck]).

Soon afterwards Aldous in [A] was able to apply the coupling arguments to the analysis of random walks on finite groups. Among several examples, he was able to find good bounds for random walk on the symmetric group generated by adjacent transposition, which do not make a conjugacy class.

A few years later Aldous and Diaconis in joint work [AD1], [AD2], introduced another kind of probabilistic argument they call the strong stationary time approach. Among several examples, they were able to analyze "top to random" shuffle which was subsequently studied and generalized by others. Using the strong uniform time approach Broder and Matthews were able to give a simple analysis of the random walk on a cube and of the random walk on a symmetric group generated by transpositions (see [Matt1], [D], §4B).

The theory of strong stationary times was later developed by Diaconis and Fill in [DF], where they introduced an important concept of dual processes. There they were able to analyze birth and death chains and introduced a criteria for perfect strong stationary times (in our terminology). Also Matthews recently found connection between strong stationary times and eigenvalues (see [Matt3]).

In recent works by Diaconis and Saloff-Coste other analytic techniques were introduced including the comparison technique (see e.g. [DSC], [D-S-C]). See [Dc] for a review of the recent developments.

We must add to this picture a tremendous amount of work done by a generation of the graduate students who were able to obtain sharp bounds in various important cases. Note the recent papers [Hild] and [Stong1] of Hilderbrand and Stong, where they analyze a random walk on $SL(n;\mathbb{F}_q)$ generated by transvections and a natural random walk on the Heisenberg group mod p.

Now let me briefly describe the strong uniform time approach and why it is useful for the study of the random walks on groups.

Suppose we are given a group G, a set of generators S and a probability distribution $\mathbf{P}$ on S. Start at the identity element $e \in G$. At each step of the walk we pick a generator from the probability distribution $\mathbf{P}$ and multiply by it on the right. Under mild conditions the above defined Markov chain is ergodic i.e. converges to a uniform distribution on G. There are various ways to quantify the rate of convergence, including the most commonly used total variation distance, these are discussed in section 2 below. Following [AD1], [AD2] we consider *separation distance*

$$s_k = 1 - |G| \cdot \min_{g \in G} Q^k(g)$$

where Q^k is the probability distribution of the state of the walk after k steps. It turns out that the sequence s_k is nonincreasing, submultiplicative and has simple relations with other mixing times (see §2.2, 2.4). As k tends to infinity the separation distance s_k becomes exponentially small:

$$s_k \sim C\, e^{-\alpha\, k}$$

(see §2.2, 3).

Suppose we are given an algorithm, some kind of *stopping rule* which stops the walk. Its stopping time τ which can be thought as a random variable, is called *strong uniform time* if the stopping state is random even conditioned on the time we stopped (see §3.1 for the precise definition). The following bound due to Aldous and Diaconis justifies this seemingly strange notion:

$$s_k \leq P(\tau > k)$$

In [AD2] they also show the existence of a strong uniform time for which the above inequality becomes an equality for all k. Such strong uniform times are called *perfect*. An easy necessary condition (see [DF]) says that if τ has a *halting state* i.e. a state such that whenever the walk hits it, τ stops there, then τ is perfect.

It is surprisingly hard to find strong uniform times. These have only been found for special classes of walks. When a hypotesized time is found, it is sometimes difficult to show that the stopping rule defines a strong uniform time. However, it is usually not very hard to estimate the expectation of the stopping time. In this thesis we introduce the notion of *total separation*:

$$s = 1 + s_1 + s_2 + \dots$$

For a strong uniform time τ we have a bound

$$s \leq E(\tau)$$

and also

$$s_k \leq \frac{E(\tau)}{k}$$

As will be shown, s is interesting by itself as a new notion of mixing time particular because of its relations to hitting times and to cover time (see §2.2, §2.4, §3.6).

The main purpose of this work is to present explicit constructions in various particular cases which enable us to estimate and sometimes to compute exactly the total separation. The idea is to show the power of the strong uniform time approach,

rather than to get tight bounds in few particular cases. In many examples (see below) we were able to analyze walks which could not be analyzed by any other existing techniques. These examples include various random walks on symmetric groups, hyperoctahedral groups, Heisenberg groups, and others.

The thesis is constructed as follows. We start with standard notation, definitions and examples about finite groups. In chapter §2 we define random walks on groups, separation distance and study various numerical and asymptotic properties of the total separation. Notably, in §2.3 we prove that $\mathbf{P} \in \mathbb{Q}^S$ implies $s \in \mathbb{Q}$. We give examples of interesting walks (§2.5), present some bounds on s in cases of rapidly mixing random walks (§2.6) and state the *Diameter Conjecture* (§2.7).

The whole of chapter 3 is dedicated to the study of various stopping times, their relations and connections to the bounds on separation distance. In §3.1 we define stopping times and establish a language for describing them. In $\S 3.2 - 3.4$ we define strong uniform times, study them and show the existence of perfect times. In §3.5 we introduce a new notion of time-invariant stopping times which turn out to be important constructing bricks for strong uniform times. In §3.6 we study hitting times in connection with strong uniform times. In particular, we find there a lower bound on s to supplement the upper bounds that can be obtained from strong uniform time arguments.

Chapter 4 compiles some important constructions of strong uniform times in traditionally studied cases. In §4.1 we compute the total separation for the random walk on a cycle (group $\mathbb{Z}_m$). In §4.2 we we show that it is possible to analyze convergence if we slow down the walk. In $\S 4.3 - 4.5$ we show that under certain conditions one can construct a strong uniform time for some direct, semidirect and skew product of groups once strong uniform times are known for the factors. By use of our technique we are able to provide an analysis in many interesting examples of walks such as a weighted walk on a cube (§4.3), a walk on the dihedral group (§4.4) and a walk on the hyperoctahedral group (§4.5).

In chapter 5 we investigate walks on a symmetric group generated by different classes of transpositions. These classes include all transpositions (§5.2), star transpositions (§5.1), adjacent transpositions (§5.3), weighted transpositions ($\S 5.4, 5$), and semi-random transpositions (§5.8). We also study there the nearest neighbor random walk on the k-subsets of an n-set (§5.7).

In chapter 6 we consider miscellaneous geometric random walks including random walks on the upper triangular matrices ($\S 6.3, 4$), random walks on the k-subspaces of an n-dimensional space (§6.2), and the affine random walk on a $\mathbb{F}_q^n$ (§6.5). In the first section §6.1 we discuss two different ways to generate a nonsingular matrix over a finite field.

Acknowledgments

I would like to thank my advisor Persi Daiconis for his thoughtful guidance, help and for giving me a feel for randomness. Without him this thesis would never be written.

Many thanks to Richard Stanley for a number of indispensable conversations and for his impeccable taste in Combinatorics. It was a great pleasure to be around MIT in the past few years.

Since High Schools many people in mathematics helped me and advised me. I would like to thank them here: Vladimir I. Arnold, Israel M. Gelfand, Alexander A. Kirillov (Sr.), Andrew Odlyzko, Grigory Olshansky, Vladimir S. Retakh and Andrei V. Zelevinsky.

During work on this thesis I had conversations with a number of people who commented my work and helped me understand the subject better. Thank you Noga Alon, Alex Astashkevich, Gene Cooperman, Peter Matthews, Agoston Pisztora, Boris Pittel, Gian-Carlo Rota, Larry Shepp, David B. Wilson and Peter Winkler. I am especially grateful for the interesting conversations with David Aldous and Jim Fill, whose works lie in the foundations of the theory of strong stationary times.

Special thanks to Sergey Bratus and Nathan Lulov who read and commented parts of the thesis. Converstations with them were very stimulating.

Finally, I would like to thank my parents, Mark and Sofia Pak for their patience, help and encouragements at all stages of my life. This thesis is dedicated to them.

Notation

$\mathbb{N} = \{1, 2, \dots\}$

$\mathbb{Z}_+ = \{0, 1, 2, \dots\}$

$[n] = \{1, 2, \dots, n\}$

$\begin{bmatrix} n \\ k \end{bmatrix}$ - k-subsets of an n-set $[n]$

$n! = 1 \cdot 2 \cdot \ldots \cdot n$, $n \in \mathbb{N}$

$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \left| \begin{bmatrix} n \\ k \end{bmatrix} \right|$

$(n)_q = 1 + q + q^2 + \cdots + q^{n-1}$, $n \in \mathbb{N}$

$(n!)_q = (1)_q \cdot (2)_q \cdot \ldots \cdot (n)_q$, $n \in \mathbb{N}$

$\binom{n}{k}_q = \frac{(n!)_q}{(k!)_q\,(n-k)!_q}$

$\mathfrak{h}_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ - partial harmonic sum

$\mathbb{Z}$ - group of integers

$\mathbb{Z}_m \simeq \mathbb{Z}/m\mathbb{Z}$ - cyclic group with m elements

$\mathbb{F}_q$ - finite field with q elements

$GL(n; \mathbb{F}_q)$ - full linear group over the field with q elements

$U(n; \mathbb{F}_q)$ - unipotent group over the field with q elements

S_n - group of permutations of n elements

A_n - alternating group of permutations of n elements

$B_n = S_n \ltimes \mathbb{Z}_2^n$ - group of symmetries of an n-dimensional cube

$\mathcal{W} = (G, S, \mathbf{P})$ - random walk on a group G with a set of generators S and probability distribution $\mathbf{P}$ on S (see §2.1)

$f(x) = O(g(x))$ as $x \to \infty$ means that $\frac{f(x)}{g(x)}$ is bounded as $x \to \infty$

$f(x) = o(g(x))$ as $x \to \infty$ means that $\frac{f(x)}{g(x)} \to 0$ as $x \to \infty$

$f(x) \sim g(x)$ as $x \to \infty$ means that $\frac{f(x)}{g(x)} \to 1$ as $x \to \infty$

$f(x) \approx C \cdot g(x)$ means that there exist $C < \infty$, $\frac{f(x)}{g(x)} \to C$ as $x \to \infty$

1. Groups

1.1 Definitions and general properties.

We will consider only finite groups unless otherwise specified. Suppose we have a finite group G. The identity element will be denoted by e. The *degree* of an element $g \in G$ is the minimal degree $m \in \mathbb{Z}_+$ such that $g^m = e$. By $|G|$ we denote the *order* of G, i.e. the number of elements in G.

We use left-to-right notation for a product of elements. This means that when we say "multiply a by b" we mean an element $a \cdot b$. By "$\simeq$" we denote an isomorphism of groups.

The *direct product* of groups G_1, G_2 is a group $G = G_1 \times G_2$ with the set of elements $\{(g_1, g_2), g_1 \in G_1, g_2 \in G_2\}$, the identity element $e = (e_1, e_2)$ and multiplication law

$$(g_1, g_2) \cdot (g_1', g_2') = (g_1 g_1', g_2 g_2')$$

(see e.g. [VDW], §53) Denote $G^k = G \times G \times \cdots \times G$ (k times).

$Aut(G)$ is the group of automorphisms of G. We say that group G_1 *acts* on G_2 if there is a homomorphism

$$f : G_1 \to Aut(G_2)$$

The *semidirect product* of G_1 acting on G_2 is a group $G = G_1 \ltimes G_2$ with the set of elements $\{(g_1, g_2), g_1 \in G_1, g_2 \in G_2\}$, identity element $e = (e_1, e_2)$ and multiplication law

$$(g_1, g_2) \cdot (g_1', g_2') = \left(g_1 g_1', g_2 \cdot f^{-1}(g_1)\,[g_2']\right)$$

(see e.g. [DM], §2.5) By definition, $|G_1 \ltimes G_2| = |G_1| \cdot |G_2|$. Note also that if f is a trivial homomorphism $f(g) = e_2$, $g \in G_1$, the semidirect product becomes a direct product.

Let $G \subset S_n$ be a *permutation group*. By the *wreath product* of groups G and H we mean a semidirect product $G \ltimes H^n$, where G acts on H^n by permuting copies of H (see e.g. [DM], §2.6).

The group G is *generated* by S, $S \subset G$ if any element $g \in G$ can be written as

$$g = s_1 \cdot s_2 \cdot \ldots \cdot s_l$$

for some $s_1, s_2, \ldots, s_l \in S$. The set S is called a *set of generators* and its elements are called *generators*. We always include the identity e in the set of generators unless explicitly stated even if we don't list it when describing the set S. If $e \in S$ denote by $\check{S}$ the set of nontrivial generators of G:

$$\check{S} = S \setminus \{e\}$$

Define the *length* $l_S(g)$ of an element $g \in G$ to be the minimal number of generators needed in order to write g as their product. The *diameter* d_S is the length of the furthest element:

$$d_S = \max_{g \in G} l_S(g)$$

The *diameter set* $D_S \subset G$ is the set of all group elements $g \in G$, such that $l_S(g) = d_S$.

A set of generators S is called *symmetric* if $s \in S$ implies $s^{-1} \in S$. A set of generators S is called *minimal* if for each $s \in \check{S}$, $S \setminus s$ is no longer a set of generators. If S is both a minimal and symmetric set of generators, each generator $s \in S$ is an *involution* $s^2 = e$.

For a finite group G and its set of generators S the *Cayley graph* $\Gamma = \Gamma(G, S)$ is a graph with elements $g \in G$ as vertices and pairs $(g, g \cdot s)$, $g \in G$, $s \in \check{S}$ as edges (see [CM], §3.1). This oriented graph becomes *simple* (unoriented) when S is symmetric.

For any subset $S_1 \subset G$ denote by $\langle S_1 \rangle$ the subgroup of G generated by S_1.

We need to introduce a few combinatorial definitions designed specifically for our purposes.

An *alphabet* is a set of generators S. Its elements we also call *letters*. A *word* is any formal product of generators

$$\omega = s_1 s_2 \dots s_l$$

where $s_i \in S$, $1 \le i \le l$. The number $l = l(\omega)$ is called the *length of a word* ω. A *language* $S^\star$ is a set of all possible words which can be made out of an alphabet. The language $S^\star$ has a semigroup structure with multiplication defined as concatenation of words. The *empty word* $(\emptyset)$ is an identity element in $S^\star$ and the only element of zero length. The *value of a word* $\omega \in S^\star$ is an element $g = \gamma(\omega) \in G$ obtained by substitution of the formal product by a group product. The map $\gamma : S^\star \to G$ is a semigroup homomorphism. By $S^\star(g)$, $g \in G$ we denote the set of all words $\omega \in S^\star$ such that $\gamma(\omega) = g$. This means that $S^\star(g)$ is a preimage of g:

$$S^\star(g) = \gamma^{-1}(g)$$

By $S^\star_l(g)$ $(S^\star_l)$ we denote the set of all words $\omega \in S^\star(g)$ $(\omega \in S^\star)$ with length $l(\omega) = l$.

Call a *subword* of a word ω the word obtained as a result of erasing some of the letters of ω followed by the concatenation of what is left. Suppose $S_1 \subset S$, $\omega \in S$. An S_1*-subword* of a word ω is the word obtained as a result of erasing all the letters $s \notin S_1$ followed by the concatenation of what is left. Denote this word $\omega|_{S_1}$.

A sequence of subsets $(A_1, A_2, \dots, A_r)$, $A_i \in G$ is a *decomposition* of G if any element $g \in G$ can be *decomposed* i.e. written as

$$g = a_1 \cdot a_2 \cdot \ldots \cdot a_r$$

where $a_i \in A_i$, $1 \le i \le r$. A decomposition is called *uniform* if each element $g \in G$ has the same number of ways to be decomposed. Obviously, this number is

$$m = |A_1| \cdot |A_2| \cdot \ldots \cdot |A_r| / |G|$$

A decomposition is called *exact* if $m = 1$ i.e. each element $g \in G$ has exactly one way to be decomposed. In this case the word

$$\omega = a_1 a_2 \dots a_r$$

is called a *normal form* of a word $g \in G$, $\omega \in S^\star(g)$. For example, the sequence $(\{0,2\},\ \{0,1\})$ is an exact decomposition of the group $\mathbb{Z}_4$ of integers modulo 4.

Let $P_1, P_2 : G \to \mathbb{Q}$ be two functions on G. By the *convolution* $P_1 * P_2$ we mean a function

$$P_1 * P_2\ (g) = \sum_{h \in G} P_1(h) \cdot P_2(h^{-1}\, g)$$

In probabilistic terms this corresponds to first picking an element according to P_1, then multiplying on the right by an element picked according to P_2.

By the *group algebra* $\mathbb{Q}[G]$ we mean the ring of linear combinations $a_1\, g_1 + \cdots + a_n\, g_n$, $a_1, \ldots, a_n \in \mathbb{Q}$, $g_1, \ldots, a_g \in G$. The algebra $\mathbb{Q}[G]$ is dual to ring of $\mathbb{Q}$-valued functions on G with convolution as a multiplication.

1.2 Examples of groups.

Let the *cyclic group* $\mathbb{Z}_m$ be the group of integers module m with addition as a group operation. Here we have $e = 0$ and $|\mathbb{Z}_m| = m$. The set $S = \{\pm 1\}$ is a symmetric set of generators. The corresponding Cayley graph is a cycle. The diameter is $d_S = \left\lfloor \frac{m}{2} \right\rfloor$ and the diameter set is $D_S = \{\pm d_S\} \subset \mathbb{Z}_m$.

The *dihedral group* DH_m is the group of symmetries of the equilateral polygon (see e.g. [CM], §1.5). It is not hard to see that $DH_m \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_m$ where an action of $\mathbb{Z}_2$ on $\mathbb{Z}_m$ is defined by a nontrivial involutive automorphism $(i) \to (-i)$, $1 \le i \le m$. We have $|DH_m| = 2\, m$.

The set $S = \{(1,0), (0, \pm 1)\}$ is a symmetric set of generators. The corresponding Cayley graph is a 1-skeleton of an m-prism. The diameter is $d_{S'} = 1 + \left\lfloor \frac{m}{2} \right\rfloor$. The diameter set is $D_{S'} = \{(1, \pm \left\lfloor \frac{m}{2} \right\rfloor)\}$.

The set $S' = \{(1,0), (1,1)\}$ is a minimal symmetric set of generators. The corresponding Cayley graph is a cycle of length $2\, m$. The diameter is $d_{S'} = m$, the diameter set is $D_{S'} = \{(0,1)\}$.

The *n-cube* the group $\mathbb{Z}_2^n$. Let $S = \{(0, \ldots, 1_i, \ldots, 0), 1 \le i \le n\}$. This is a symmetric minimal set of generators. The corresponding Cayley graph is a 1-skeleton of the n-dimensional cube. We have $d_S = n$ and $D_S = \{(1, \ldots, 1)\}$.

Group 21 is a unique group of the order 21 that is not isomorphic to $\mathbb{Z}_{21}$. Group 21 can be presented as a semidirect product $\mathbb{Z}_3 \ltimes \mathbb{Z}_7$ with a an action of $\mathbb{Z}_3$ on $\mathbb{Z}_7$ defined by a nontrivial automorphism of order 3:

$$\sigma = (0, 2, 4, 6, 1, 3, 5) : \mathbb{Z}_7 \to \mathbb{Z}_7$$

In general, if p, q are primes with $q = 1 \mod (p)$ there is a unique non abelian group $G \simeq \mathbb{Z}_p \ltimes \mathbb{Z}_q$ of order $p\, q$ (see e.g. [Hall]).

The *symmetric group* S_n is the group of all permutations of an n-set $[n] = \{1, 2, \ldots, n\}$. We use the standard notation to denote permutations: $\sigma = (3,1,2) \in S_3$ means that $\sigma(1) = 3$, $\sigma(2) = 1$ and $\sigma(3) = 2$. We have $e = (1, 2, \ldots, n)$ and $|S_n| = n!$. We say that i is a *fixed point* of a permutation σ if $\sigma(i) = i$.

A permutation σ is a *k-cycle* $(i_1, i_2, \ldots, i_k)$ if $\sigma(i_1) = i_2$, $\sigma(i_2) = i_3$, $\ldots$, $\sigma(i_k) = i_1$, and $\sigma(j) = j$ otherwise. A *transposition* is a 2-cycle (i, j). In other words, transpositions are permutations with exactly $n - 2$ fixed points. An *adjacent transposition* is a transposition $(i, i+1)$, $1 \le i \le n-1$. The set S of adjacent

transpositions is a symmetric minimal set of generators of S_n. Define the *length* $l(\sigma)$ of a permutation to be its length in terms of this particular generating set: $l(\sigma) = l_S(\sigma)$. We have

$$\sum_{\sigma \in S_n} q^{l(\sigma)} = (n!)_q$$

(see e.g. [Stanley], §1.3.10) The corresponding Cayley graph Γ is called the *weak Bruhat order of* S_n (see [Stan], §$Ex.$ 3.75). It is known that Γ is a 1-skeleton of a permutohedra which can be defined as a convex hull of the points $(\sigma(1), \ldots, \sigma(n)) \in \mathbb{R}^n$, $\sigma \in S_n$ (see e.g. [EKK], §5.3, [Z], Ex. 0.10). In case $n = 3$, Γ is a 6-cycle.

Note that if the set S of generators of S_n consists of all transpositions, then $d_S = n - 1$ and D_S is the set of all n-cycles.

A permutation $\sigma \in S_n$ is called *odd* (*even*) if its length $l(\sigma)$ is odd (even). The *Alternating group* A_n is the subgroup of S_n of even permutations. We have $|A_n| = n!/2$. The set of 3-cycles is a symmetric set of generators of A_n, $n \geq 3$ (see [CM], §6.3).

Define an action of $\mathbb{Z}_3$ on $\mathbb{Z}_2^2$ by a cyclic permutation of the three nonidentical elements (note that the product of either two of them is the third one so this action is correctly defined). Observe that $A_4 \simeq \mathbb{Z}_3 \ltimes \mathbb{Z}_2^2$ (see [Serre], §9.3).

The *hyperoctahedral group* B_n is the symmetry group of the hyperoctahedron with vertices $(0, 0, \ldots, \pm 1, \ldots, 0) \in \mathbb{R}^n$ (see [CM], §7.3). In our notations, $\varpi = (3, -1, 2)$ means that we have a linear transformation $\varpi : \mathbb{R}^3 \to \mathbb{R}^3$, such that $\varpi(x_1, x_2, x_3) = (x_3, -x_1, x_2)$, $n = 3$. We have $e = (1, 2, \ldots, n)$, $|B_n| = n! \cdot 2^n$. The subgroup of B_n with no minus signs inside the brackets is tautologically isomorphic to S_n. Denote $\theta_i = (1, 2, \ldots, -i, \ldots, n)$. Then θ_1 and the set of adjacent transpositions in S_n make a symmetric minimal set of generators of B_n. The corresponding Cayley graph is called the *weak Bruhat order of* B_n. It is a 1-skeleton of an n-dimensional convex polytope which can be defined as the convex hull of the points $(\pm\sigma(1), \ldots, \pm\sigma(n)) \in \mathbb{R}^n$, $\sigma \in S_n$.

Note that B_n is a wreath product of S_n and $\mathbb{Z}_2$. In other words,

$$B_n \simeq S_n \ltimes \mathbb{Z}_2^n$$

where S_n acts on $\mathbb{Z}_2^n$ by the permutation of coordinates.

Note that $B_2 \simeq DH_4$ since both groups are defined as the symmetry group of a square. The map $\phi : B_2 \to DH_4$ defined by $\phi(2, 1) = (1, 1)$, $\phi(\theta_1) = \phi(-1, 2) = (1, 0)$ establishes an isomorphism. Therefore the Cayley graph in this case is an 8-cycle.

For any $q = p^d$, p - prime, there is a unique field $\mathbb{F}_q$ with q elements (see e.g. [VDW], §43; [Lang], §7.5). By $\mathbb{F}_q^*$ we denote its multiplicative group $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$. It is known (see e.g. [Lang], §7.5) that

$$\mathbb{F}_q^* \simeq \mathbb{Z}_{q-1}$$

The *full linear group* $GL(n; \mathbb{F}_q)$ is the group of nonsingular matrices over the finite field $\mathbb{F}_q$ with q elements. We have

$$|GL(n; \mathbb{F}_q)| = (q^n - 1) \cdot (q^n - q) \cdot \ldots \cdot (q^n - q^{n-1}) = q^{\binom{n}{2}} (q-1)^n (n!)_q$$

(see e.g. [CM], §7.4)

The *special linear group* $SL(n;\mathbb{F}_q)$ is the group of matrices $M \in Gl(n;\mathbb{F}_q)$ with determinant one:

$$det(M) = 1$$

Denote $PSL(n;\mathbb{F}_q)$ the factor of $SL(n;\mathbb{F}_q)$ over the scalars:

$$PSL(n;\mathbb{F}_q) \simeq SL(n;\mathbb{F}_q)/(\lambda \cdot Id)$$

where $\lambda \in \mathbb{F}_q$, $\lambda^n = 1$, and Id is an identity matrix (see e.g. [CM], §7.4).

The *Borel group* $B(n;\mathbb{F}_q)$ is the group of upper triangular matrices over the finite field $\mathbb{F}_q$. We have $B(n;\mathbb{F}_q) \subset GL(n;\mathbb{F}_q)$, $|B(n;\mathbb{F}_q)| = q^{\binom{n}{2}}(q-1)^n$.

The *unipotent group* $U(n;\mathbb{F}_q)$ is the group of upper triangular matrices over the finite field $\mathbb{F}_q$ with ones on diagonal. We have $U(n;\mathbb{F}_q) \subset B(n;\mathbb{F}_q)$, $|U(n;\mathbb{F}_q)| = q^{\binom{n}{2}}$.

The *Heisenberg group* mod p, p is a prime, is the group $U(3;\mathbb{F}_p)$. Observe that $U(3,\mathbb{F}_p) \simeq \mathbb{Z}_p \ltimes \mathbb{Z}_p^2$ with the action defined by a nontrivial automorphism $(x,y) \to (x, x+y)$ of order p. It is not hard to show that $U(3;\mathbb{F}_p)$ is generated by matrices

$$R_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad R_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

In case $p = 2$ the set $S = \{R_1, R_2\}$ is a minimal symmetric set of generators. Recall that $U(3,\mathbb{F}_2) \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_2^2 \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_4$. Therefore $U(3,\mathbb{F}_2) \simeq DH_4$ with an isomorphism given by a map $\phi : U(3,\mathbb{F}_2) \to DH_4$, $\phi(R_1) = (1,0)$, $\phi(R_2) = (1,1)$. Therefore the Cayley graph in this case is an 8-cycle.

1.3 Classes of groups.

A group G is called *abelian* if any two elements $g_1, g_2 \in G$ *commute*:

$$g_1 g_2 = g_2 g_1$$

Both the cyclic group and the n-cube are abelian groups. Every abelian group is a direct product of cyclic groups (see e.g. [VDW], §53). Denote $\mathcal{A}b$ a family of abelian groups.

A *tower of subgroups* is a sequence of the following type:

$$(*) \qquad G = G_0 \supset G_1 \supset G_2 \supset \ldots \supset G_n \supset G_{n+1} = \{e\}$$

A tower of subgroups is called *subnormal* if for each $i = 1, \ldots, n$, G_i is a normal subgroup of G_{i-1} (see e.g. [VDW], §51). By the *factors* in a subnormal tower we mean the factor groups

$$H_i = G_{i-1}/G_i$$

where $1 \le i \le n+1$.

A group $G \neq \{e\}$ is called *simple* if it has only two normal subgroups: G and $\{e\}$ (see e.g. [VDW], §51). In other words it has only the *trivial* subnormal tower: $G \supset \{e\}$.

A subnormal tower without repetitions which cannot be refined is called a *composition series*. It is not hard to see that in a composition series all factors H_i are simple. Two composition series are called *equivalent* if the factors of the first series are isomorphic to the factors of the second series, up to some permutation of indices. The Jordan-Hölder Theorem states that all composition series of a finite group are equivalent (see e.g. [VDW], §51).

There is a classification theorem for the finite simple groups. Even stating it is a nontrivial task. It basically says that every simple groups is either $\mathbb{Z}_p$, p - prime, or A_n, $n \geq 5$, or a group of Lie type, or one of the 26 sporadic groups (see [Gor] for definitions and details). By a group of Lie type we mean a member of the few known families of the finite linear algebraic groups such as $PSL(n, \mathbb{F}_q)$ and its analogs for the other root systems.

A group G is called *solvable* if it has a normal tower $(*)$ such that all factors $H_i = G_{i-1}/G_i$ are abelian, $i = 1, \ldots, n+1$ (see e.g. [Serre], §9.3).

A group G is called *supersolvable* if it has a tower $(*)$ such that G_i is the a normal subgroup of G and the factors $H_i = G_{i-1}/G_i$ are cyclic, $i = 1, \ldots, n$ (see e.g. [Serre], §9.3).

A group G is called *nilpotent* if if it has a normal tower $(*)$ such that $(G_i, G) \subset G_{i+1}$ for each $i = 0, 1, \ldots, n$ (see e.g. [Serre], §9.3).

Denote $\mathcal{S}olv$, $\mathcal{S}uperSolv$, $\mathcal{N}il$ the families of solvable, supersolvable and nilpotent groups respectively. We have

$$\mathcal{A}b \subset \mathcal{N}il \subset \mathcal{S}uperSolv \subset \mathcal{S}olv$$

with all the inclusions strict as we show later (see e.g. [Serre], §9.3).

A group G is called *p-group* if $|G| = p^k$, p - prime, $k > 0$. Every p-group is nilpotent (see e.g. [Serre], §9.3, [Lang], §1.6).

The first nonabelian group is S_3. Its only nontrivial normal tower is

$$S_3 \supset A_3 \supset \{e\}$$

This shows that S_3 is solvable, even supersolvable, but not nilpotent.

The *Klein 4-group* $K \subset A_4$ is the group of four permutations $(1,2,3,4)$, $(2,1,4,3)$, $(3,4,1,2)$ and $(4,3,2,1)$. K is the only nontrivial normal subgroup of A_4, $K \simeq \mathbb{Z}_2^2$. Therefore A_4 has a unique nontrivial normal tower:

$$A_4 \supset K \supset \{e\}$$

This shows that A_4 is solvable but not supersolvable.

As noted before, the groups A_n, $n \geq 5$ are simple. The symmetric group S_n for $n \geq 5$ contains exactly one normal subgroup A_n and therefore has the unique nontrivial normal tower:

$$S_n \supset A_n \supset \{e\}$$

Therefore the groups A_n, S_n, $n \geq 5$ are not solvable.

Group 21 is the smallest nonabelian group of the odd order. By the Feit-Thompson Theorem every finite group of odd order is solvable (see e.g. [Gor], §1).

Denote by $U(n,k;\mathbb{F}_q) \subset U(n;\mathbb{F}_q)$, $1 \leq k \leq n-1$ a subgroup of unipotent matrices $(u_{i,j})$ such that $u_{i,j} = 0$ if $1 \leq j - i \leq k$. We have a tower of normal subgroups

$$U(n;\mathbb{F}_q) \supset U(n,1;\mathbb{F}_q) \supset U(n,2;\mathbb{F}_q) \supset \ldots \supset U(n,n-1;\mathbb{F}_q) = \{e\}$$

which shows that $U(n;\mathbb{F}_q)$ is nilpotent.

Observe that $U(3;\mathbb{F}_2)$ is nilpotent but not abelian. We have $|U(3;\mathbb{F}_2)| = 8$, which makes it the smallest group with such property since the only smaller nonabelian group is S_3 (see e.g. [CM], Table 1). Recall that $U(3;\mathbb{F}_2) \simeq B_2 \simeq DH_4$.

Recall that $B(n;\mathbb{F}_q) \supset U(n;\mathbb{F}_q)$. Adding the above tower of normal subgroups of $U(n;\mathbb{F}_q)$ we get a new tower which shows that $B(n,\mathbb{F}_q)$ is solvable. When q is a prime, $B(n;\mathbb{F}_q)$ is also supersolvable. Since some of the factors may not be cyclic one needs to refine this tower. We will not use this refinement. It is not hard to show that $B(n,\mathbb{F}_q)$ is not nilpotent.

It is known that $PSL(n;\mathbb{F}_q)$ is simple for $n \geq 3$, and for $n = 2$, $q \geq 4$. Therefore neither $GL(n;\mathbb{F}_q)$ nor $SL(n;\mathbb{F}_q)$ is solvable.

2. Random walks

2.1 Definitions.

Fix a finite group G and a set of generators S. Let $\mathbf{P}$ be a probability distribution on S. A *random walk* $\mathcal{W} = (G, S, \mathbf{P})$ is a discrete Markov chain X_t (see e.g. [Feller], §15.1) with the set of states G which starts at $X_0 = e$ and moves with transition probabilities

- $P(g \to g \cdot s) = \mathbf{p}(s)$, $s \in S$, $g \in G$
- $P(g \to g') = 0$ otherwise, $g, g' \in G$

Call a probability distribution $\mathbf{P}$ on a set of generators S *uniform* if $\mathbf{p}(s_1) = \mathbf{p}(s_2)$ for all $s_1, s_2 \in \check{S}$. In this case we can think of our random walk $\mathcal{W}$ as the nearest neighbor random walk on a Cayley graph $\Gamma = \Gamma(G, S)$.

Two random walks $\mathcal{W}_1 = (G_1, S_1, \mathbf{P}_1)$ and $\mathcal{W}_1 = (G_2, S_2, \mathbf{P}_2)$ are called *isomorphic* (denoted as $\mathcal{W}_1 \cong \mathcal{W}_2$) if there is a group isomorphism $\phi : G_1 \to G_2$ such that $\phi(S_1) = S_2$ and $\phi(\mathbf{P}_1) = \mathbf{P}_2$, which means that $s_1 \in S_1$ implies $s_2 = \phi(s_1) \in S_2$ and $\mathbf{p}_1(s_1) = \mathbf{p}_2(s_2)$.

Denote by $Q^k(g) = P(X_k = g)$, $g \in G$ the probability distribution of the state of a walk after k steps. Obviously $Q^k(g) = \mathbf{P} * \mathbf{P} * \cdots * \mathbf{P}$ (k times). For a subset $S \subset G$ the probability distribution $R(g)$ is called *uniform on* S, denoted $R = U_S$, if $R(g) = \frac{1}{|S|}$ if $g \in S$ and $R(g) = 0$ otherwise.

Two random walks $\mathcal{W}_1 = (G_1, S_1, \mathbf{P}_1)$ and $\mathcal{W}_1 = (G_2, S_2, \mathbf{P}_2)$ are called *equivalent* (denoted as $\mathcal{W}_1 \simeq \mathcal{W}_2$) if there is a one-to-one map $\phi : G_1 \to G_2$ which maps probability distribution Q_1^k into Q_2^k for all $k > 0$. In other words,

$$Q_1^k(g) = Q_2^k(\phi(g)), \ g \in G_1, k > 0$$

Note that if $\mathcal{W}_1 \simeq \mathcal{W}_2$, $G_1 \simeq G_2$, $\phi(S_1) = S_2$, we immediately get

$$\mathbf{p}_1(s) = Q_1^1(s) = Q_2^1(\phi(s)) = \mathbf{p}_2(\phi(s)), s \in S_1$$

i.e. in this case $\mathcal{W}_1 \cong \mathcal{W}_2$.

Example 2.1.1 The first example to consider is the random walk on $\mathbb{Z}_m$. Take a set of generators $S = \{-1, 0, +1\}$ and a probability distribution $\mathbf{p}(0) = \frac{1}{2}$ and $\mathbf{p}(-1) = \mathbf{p}(+1) = \frac{1}{4}$. One can think of this random walk as a random walk on a circle with probability of staying to be $\frac{1}{2}$ and equal probability of moving in each direction.

Example 2.1.2 Denote by $\mathcal{W}_a$ a random walk on $\mathbb{Z}_m$ with the set of generators $S = \{0, \pm a\}$ and a probability distribution $\mathbf{p}(0) = \frac{1}{2}$ and $\mathbf{p}(\pm a) = \frac{1}{4}$. It is not hard to see that $\mathcal{W}_a \cong \mathcal{W}_1$ for all a, $1 \le a \le m$, $(a, m) = 1$.

Example 2.1.3 Consider the Heisenberg group mod 2, i.e. $U(3, \mathbb{F}_2)$ generated by matrices R_1 and R_2 (see §1.2). Take a probability distribution $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(R_1) = \mathbf{p}(R_2) = \frac{1}{4}$. Denote by $\mathcal{W}$ a uniform random walk $(U(3, \mathbb{F}_2), S, \mathbf{P})$, where $S = \{R_1, R_2\}$. Since the Cayley graph $\Gamma(U(3, \mathbb{F}_2), S)$ is an 8-cycle, $\mathcal{W}$ is equivalent to the random walk on $\mathbb{Z}_8$ although not isomorphic to it.

It is natural to assume that our probability distribution $\mathbf{P}$ is strictly positive on S i.e. for all $s \in S$, $\mathbf{p}(s) > 0$. $\mathbf{P}$ is called *symmetric* if S is symmetric and $\mathbf{p}(s^{-1}) = \mathbf{p}(s)$ for all $s \in S$. In the future we restrict our attention only to random walks $\mathcal{W} = (G, S, \mathbf{P})$ where G is a finite group, S is a symmetric set of generators, $e \in S$, and $\mathbf{P}$ is symmetric and strictly positive on S. We call such random walks *directed*.

A classical result of the Markov Chain Theory is the following theorem (see e.g. [AF], §2.1).

Theorem 2.1.4 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk. Then the probability distribution $Q^k(g)$ tends to uniform stationary distribution $\pi(g) = \frac{1}{|G|}$ as k tends to infinity.

There is a way to define random walks in terms of words in the alphabet S. First we need to explain a few useful expressions.

When saying *draw a generator $s \in S$ from probability distribution* $\mathbf{P}$ we mean choosing a generator s with probability $\mathbf{p}(s)$. By *applying* a generator s to a word ω we mean multiplying it to the right: $\omega \cdot s$.

Consider the following procedure

Procedure W Suppose we are given a word $\omega_t \in S_t^\star$.

- Draw a generator $s_{t+1} \in S$ from the probability distribution $\mathbf{P}$
- Apply it to ω: $\omega_{t+1} \leftarrow \omega_t \cdot s_{t+1}$.
- $t \leftarrow t + 1$.

Now we can think of the random walk $\mathcal{W}$ via the use of the Procedure $\mathbf{W}$ where the state X_t is given by the value of a word ω_t:

$$X_t = \gamma(\omega_t)$$

We say that a walk X_t *walked along the word* $\omega_n \in S_n^\star$ if

$$\omega_t = s_1 s_2 \dots s_n$$

where generators $s_i \in S$, $1 \le i \le n$ are given by the formula $s_i = X_{i-1}^{-1} \cdot X_i$.

At last, we would like to give a simple combinatorial interpretation for the value $Q^k(g)$ in case $\mathbf{P}$ is uniform. Denote $p = \mathbf{p}(e)$, $q = \mathbf{p}(s)$, $s \in \check{S}$. We have $q = (1-p)/|\check{S}|$.

Theorem 2.1.5

$$Q^k(g) = \sum_{l=l_S(g)}^{k} r_i(g)\, p^{k-l}\, q^l \binom{k}{l}$$

in particular

$$Q^k(g) = r_k(g)\, q^k,\ 1 \le k \le l_S(g)$$

where $r_l(g)$ is one of the following:

i) the number of paths with length l in the Cayley graph Γ going from e to g

ii) $r_l = |S^\star_l(g)|$ i.e. the number of words ω such that $\gamma(\omega) = g$, $l(\omega) = l$.

Proof Denote by j the number of times we stayed before reaching $g \in G$ along the path ζ. The number of ways to choose the places along ζ where to stay is the binomial coefficient $\binom{k}{j}$. Therefore the probability of reaching g from e along ζ is exactly

$$p^j \, q^{k-j} \binom{k}{j}$$

Summing over all j we get part *i)*.

The proof of part *ii)* is analogous. □

2.2 Separation distance.

We would like to measure how fast our walk is mixing. In order to do that we need to define a distance on a probability space so we can bound the difference $(Q^k - U_G)$.

Definition 2.2.1 *Separation distance* s_k (see [AD1, AD2]) is defined by the formula

$$s_k = |G| \cdot \max_{g \in G} \left(\frac{1}{|G|} - Q^k(g) \right)$$

Although this separation distance from the uniform distribution is not a distance in a traditional sense, it fits our purposes since by the Theorem 2.1.4 we get a uniform stationary distribution for all walks we consider. Other distances are discussed in §2.4.

Here is another way to think of a separation distance.

Definition 2.2.2 Call an element $\widehat{g}_k$ *k-minimal* if the probability distribution Q^k on a group G has a minimum value at $\widehat{g}_k$. Define a *minimal sequence* to be any sequence $(\widehat{g}_1, \widehat{g}_2, \dots)$ of k-minimal elements for each $k > 0$.

Now suppose we have a minimal sequence $(\widehat{g}_1, \widehat{g}_2, \dots)$. By definition of the separation distance we have

$$s_k = 1 - |G| \cdot Q^k(\widehat{g}_k)$$

Note that for some k there might be many k-minimal elements. Thus a minimal sequence also is not uniquely determined, for example any $g \in G$ with $Q^k = 0$ is k-minimal.

Our goal is to give some upper bounds for the asymptotic behavior of the sequence s_k when k tends to infinity. We say that a random walk $\mathcal{W}$ is *degenerate* if $s_k = 0$ for some k.

Theorem 2.2.3 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed nondegenerate random walk. Then there exist constants $C, \alpha > 0$ such that

$$s_k \sim C \cdot e^{-\alpha \, k}$$

as k tends to infinity.

A proof of Theorem 2.2.3 will be given in §2.3. The key result for proving this theorem is the following proposition, which is also interesting by itself.

Proposition 2.2.4 The separation distances s_k, $k > 0$ satisfies the following properties

(monotonicity) $s_m \leq s_n$, $m \leq n$

(submultiplicativity) $s_m \cdot s_n \geq s_{m+n}$

A proof of the Proposition 2.2.4 in greater generality is given in [AD2] (see also [D], §4*C*).

Lemma 2.2.5 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk. Then there exist constants $C, \alpha > 0$ such that

$$s_k \leq C \cdot e^{-\alpha\, k}$$

for all $k > 0$.

Note that Theorem 2.2.3 doesn't immediately follow from the Lemma since a priori the separation distance s_k could have a different kind of asymptotic behavior, such as decrease superexponentially.

Proof of Lemma 2.2.5 Let $d = d_S$ be a diameter of the group G. By definition of a separation distance we have

$$s_1 = s_2 = \cdots = s_{d-1} = 1,\ s_d < 1$$

Now from monotonicity and submultiplicativity we get

$$s_{m\,d+i} \leq s_d^m, \quad m \geq 0, \quad 0 \leq i \leq d-1$$

Therefore

$$s_k \leq s_d^{\lfloor \frac{k}{d} \rfloor}$$

This proves the Lemma. □

Remark 2.2.6 The Definition 2.2.1 and the Proposition 2.2.4 are due to Aldous and Diaconis (see [AD1, AD2]). For generalizations and subsequent treatment see [AF, D, DF]. k-minimal elements were introduced in [DF] (see also [AF], §9.2, where they are called halting states).

Sometimes it is possible to find very tight bounds on a separation distance. We need the following definitions.

Definition 2.2.7 The *separation series* and the *total separation* are defined by the following formulas

$$s(x) = 1 + \sum_{k=1}^{\infty} s_k x^k$$

$$s = s(1) = 1 + s_1 + s_2 + \dots$$

Theorem 2.2.8 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk, d_S the diameter of G. Then

1) $d_S \le s < \infty$

2) $1 - \frac{k}{s} \le s_k \le \frac{s - d_S}{k - d_S + 1}, \ \ k \ge d_S$

Proof The first inequality in part *1)* follows immediately from the observation that $s_1 = s_2 = \cdots = s_{d_S - 1} = 1$.

The second inequality in part *1)* of the Theorem follows from Lemma 2.2.5. Indeed, we have

$$s \le 1 + C\, e^{-\alpha} + C\, e^{-2\,\alpha} + \cdots = 1 + C\, \frac{e^{-\alpha}}{1 - e^{-\alpha}} < \infty$$

where the last inequality comes from $\alpha > 0$.

From monotonicity (see Proposition 2.2.4) we have

$$s \ge d_S + s_{d_S} + \cdots + s_k \ge d_S + (k - d_S + 1)\, s_k$$

which proves the second inequality in part *2)* of the Theorem.

Analogously from monotonicity and submultiplicativity we have

$$s \le k + k\, s_k + k\, s_k^2 + \cdots = \frac{k}{1 - s_k}$$

Thus $s_k \ge 1 - \frac{k}{s}$ and we have proved the first inequality in part *2)* of the Theorem. □

For every $\epsilon > 0$ define $n_\epsilon = n_\epsilon(\mathcal{W})$ to be the smallest number such that $s_{n_\epsilon} \le \epsilon$. Heuristically, after $n_{\frac{1}{2}}$ steps the separation distance is $\frac{1}{2}$ and after that decreasing exponentially fast. Many authors consider $n_{\frac{1}{2}}$ to be a definition of the *mixing time* (see e.g. [Sinclair]).

Corollary 2.2.9 Let s be the total separation of a directed random walk $\mathcal{W} = (G, S, \mathbf{P})$. Then

$$\frac{s}{2} \le n_{\frac{1}{2}} \le 2\, s$$

Proof Take $k_1 = \left\lfloor \frac{s}{2} \right\rfloor$ in part *2)* of the Theorem 2.2.8 for the lower bound and $k_2 = \lfloor 2\, s \rfloor$ for the upper bound. □

Observe that by part *2)* of Theorem 2.2.8 we can bound the separation distances s_k, $k \ge 1$ in terms of the total separation s. This appears to be a very useful bound although it is not tight at all as n tends to infinity. Here we get $s_k \le \frac{C_1}{k}$ instead of the $s_k \le C_2\, e^{-\alpha\, k}$.

It is easy to see that the α in the Theorem 2.2.3 is a radius of convergence of $s(x)$ around zero. In this thesis rather than estimating the separation distances s_k for all k we will be finding bounds for this radius ρ and the total separation s.

Definition 2.2.10 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a random walk. Call an element $\widehat{g} \in G$ *extremal* if it is k-minimal for all $k > 0$. A random walk $\mathcal{W}$ is called *guileless* if it has an extremal element

Observe that if $\widehat{g}$ is an extremal element we have

$$s = \sum_{k=1}^{\infty} 1 - |G| \cdot Q^k(\widehat{g})$$

Theorem 2.2.11 Let $\mathcal{W}_1 \simeq \mathcal{W}_2$ be two equivalent random walks, s^1, s^2 - their total separations. Then

$$s^1 = s^2$$

Proof Obvious. □

Example 2.2.12 Consider a random walk on $G \simeq \mathbb{Z}_2$ defined as follows. Let $G = \{e, a\}$, $\mathbf{p}(e) = \frac{1}{2} + \frac{p}{2}$, $\mathbf{p}(a) = \frac{1}{2} - \frac{p}{2}$, $1 > p \geq 0$. After k steps we get a probability distribution

$$Q^k(e) = \frac{1}{2} + \frac{1}{2} \cdot p^k$$
$$Q^k(a) = \frac{1}{2} - \frac{1}{2} \cdot p^k$$

Therefore the separation distance is decreasing exponentially:

$$s_k = p^k$$

From here we get the separation series, radius of convergence and the total separation:

$$s(x) = \frac{1}{1 - p \cdot x}$$
$$s = s(1) = \frac{1}{1 - p}$$
$$\rho = \frac{1}{p}$$

Observe that $\mathcal{W}$ has an extremal element $\widehat{g} = a$. Note that if $-1 < p < 0$ we still can define a random walk $\mathcal{W}$. Clearly this random walk does not have an extreme element.

Example 2.2.13 Consider a random walk on $\mathbb{Z}_4$ with generating set $\{0, \pm 1\}$ (see Example 2.1.1). For any $k \geq 1$ we have

$$Q^k(0) = \frac{1}{4} + \frac{1}{2^{k+1}}$$
$$Q^k(2) = \frac{1}{4} - \frac{1}{2^{k+1}}$$
$$Q^k(\pm 1) = \frac{1}{4}$$

Therefore we get the exact value of the separation distance:

$$s_k = \frac{1}{2^{k-1}}$$

From here we can compute the separation series:

$$s(t) = 1 + t(1 + \frac{t}{2} + \frac{t^2}{2^2} + \dots) = \frac{2+t}{2-t}$$

which gives us the total separation and the radius of convergence:

$$s = s(1) = 3, \qquad \rho = 2$$

Example 2.2.14 Let G be a finite group, $|G| \geq 4$, $|G|$ even. Denote $N = |G|-1$. Fix an element $a \in G$, $a \neq e$. Suppose a is an involution, i.e. $a^2 = e$. Then the set $S = G \backslash \{a\}$ is a symmetric set of generators. Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$ with probability distribution $\mathbf{p}(s) = 1/N$, $s \in S$. We have

$$Q^2(g) = \frac{N-1}{N^2}, \ g \neq e$$
$$Q^2(e) = \frac{1}{N}$$

Indeed,

$$S_2^\star(g) = \{\omega = h_1\, h_2 \mid h_2 = h_1^{-1} \cdot g, \ h_1 \neq e, a\}$$

for any $g \in G$. Therefore $r_2(g) = |G|-2 = N-1$ if $g \neq e$, and $r_2(e) = |G|-1 = N$. This gives the above formulas. Observe that

$$Q^2 = \frac{1}{N^2} \cdot I + \frac{N^2-1}{N^2} \cdot J$$

where I is an identity distribution $I(e) = 1$, $I(g) = 0$, $g \neq e$, and J is a uniform distribution $J(h) = 1/|G|$, $h \in G$. From here we have

$$Q^{2\,k} = \frac{1}{N^{2\,k}} \cdot I + \left(1 - \frac{1}{N^{2\,k}}\right) \cdot J$$
$$Q^{2\,k+1} = \frac{1}{N^{2\,k}} \cdot Q^1 + \left(1 - \frac{1}{N^{2\,k}}\right) \cdot J$$

This proves that a is an extremal element, i.e. $s_k = 1-(N+1)Q^k(a)$. Therefore the separation distance is decreasing exponentially

$$s_{2\,k} = s_{2\,k+1} = \frac{1}{N^{2\,k}}, \quad k > 0$$

Thus the separation series is equal to

$$s(x) = 1 + x + \frac{x^2}{N^2} + \frac{x^3}{N^2} + \frac{x^4}{N^4} + \frac{x^5}{N^4} + \dots = N^2\,\frac{1+x}{N^2-x^2}$$

From here we get the total separation and the separation radius:

$$s = s(1) = \frac{2\,N^2}{N^2-1}, \quad \rho = N$$

2.3 Properties of the separation series.

Let $\{f_n\}$, $n \geq 0$ be a sequence of functions on the group G: $f_n : G \to \mathbb{R}$. Suppose sequence f_n tends to U_G as n tends to infinity. In other words for each $g \in G$ we have

$$\lim_{n\to\infty} f_n(g) = \frac{1}{|G|}$$

We say that $\{f_n\}$ is *asymptotically stable* if there is a number $n_0 \in \mathbb{N}$ and linear order "$\prec$" on G (called an *asymptotic order*) such that $f_n(g_1) \leq f_n(g_2)$ for all $g_1 \prec g_2$ and $n \geq n_0$. The minimal element in the asymptotic order is called *asymptotically extremal*.

Random walk $\mathcal{W} = (G, S, \mathbf{P})$ is called *lazy* if $\mathbf{p}(e) \geq \frac{1}{2}$. Most of the random walks we are going to study are lazy. There are several technical reasons to do so. Here is one of them.

Theorem 2.3.1 Let $\mathcal{W}$ be a directed random walk. Then the sequences $\{Q^{2n}\}$ and $\{Q^{2n+1}\}$ are asymptotically stable. Moreover, if $\mathcal{W}$ is lazy, the sequence $\{Q^n\}$ is asymptotically stable.

Proof Let $N = |G|$, $G = \{g_1 = e, g_2, \dots, g_N\}$. Denote by $A = \{a_{i,j}\}$ an $N \times N$ - matrix such that $a_{i,j} = Q^1(g_i^{-1} g_j)$ i.e. $a_{i,j} = \mathbf{p}(s)$ if $s = g_i^{-1} g_j \in S$ and $a_{i,j} = 0$ otherwise. Denote also $b_0 = (1, 0, \dots, 0)$ a column N-vector, $b_k = (Q^k(g_1), Q^k(g_2), \dots, Q^k(g_N))$. We have $b_k = A^k\, b_0$.

Recall that $\mathbf{P}$ is a symmetric probability distribution and $\langle S\rangle = G$, $e \in S$ and $\mathbf{p}(s) > 0$, $s \in S$. It is not hard to argue that in an appropriate basis $v_1, \dots, v_N$, $v_i = (v_{i,1}, \dots, v_{i,N})$ we have

$$b_k = \lambda_1^k \cdot v_1 + \dots + \lambda_N^k \cdot v_N$$

where $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_N > -1$ (about spectra of graphs including Cayley graphs see e.g. [CDS], [Chung], §1.3).

Suppose now $\mathbf{p}(e) \geq \frac{1}{2}$. From here we get $A = \frac{1}{2}(A_1 + id_N)$ for some matrix A_1 of the same type. Therefore $\lambda_N \geq \frac{1}{2}(-1 + 1) = 0$

Consider elements g_i and g_j. Let r be the smallest number such that

$$v_{r,i} + v_{r+1,i} + \dots + v_{l,i} \neq v_{r,j} + v_{r+1,j} \dots + v_{l,j}$$

where $\lambda_{r-1} > \lambda_r = \lambda_{r+1} = \dots = \lambda_l > \lambda_{l+1}$. We have

$$Q^k(g_i) - Q^k(g_j) = \lambda_r^k(v_{r,i} + \dots + v_{l,i} - v_{r,j} - \dots - v_{l,j}) + \lambda_{l+1}^k\, c_1 + \lambda_{l+2}\, c_2 + \dots$$

Since $\lambda_l > \lambda_{l+1} \geq \lambda_{l+2} \geq \dots$ we immediately get an asymptotic order on g_i and g_j for all $k \geq n(i,j)$. Taking n_0 to be a maximum of $n(i,j)$ over all pairs $1 \leq i, j \leq N$ we get an asymptotic order on G. This proves the second part of the Theorem.

The first part is analogous, but in this case one needs to take the matrix A^2 which also has only nonnegative eigenvalues. $\square$

For the next Theorem think of the values $p_g = \mathbf{p}(g)$, $g \in S$ as commuting variables.

Theorem 2.3.2 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk. Then

$$s(x) \in \mathbb{Z}(x;\ p_g,\ g \in S)$$

i.e. the separation series $s(x)$ is a ratio of two integer polynomials in x and p_g, $g \in S$.

Proof Let $\widehat{g}_1$ and $\widehat{g}_2$ be the asymptotically extremal elements for the sequences Q^{2k} and Q^{2k+1}, $k > 0$. By definition of the separation series we have

$$s(z) = \frac{1}{1-z} - |G| \sum_{k=1}^{\infty} \min_{g \in G} Q^k(g) z^k$$

Indeed, by Lemma 2.2.5 the series $s(z)$ is convergent at 1. Therefore $s(z)$ is also convergent at z for all $z \in \mathbb{C}$, $|z| < 1$. This implies that $s(z)$ is an analytic function in some disc with center at 0 the radius of convergence $\rho \geq 1$.

From Theorem 2.3.1 we have

$$s(z) = \frac{1}{1-z} - P(z) - |G| \sum_{k=1}^{\infty} Q^{2k}(\widehat{g}_1) z^{2k} - |G| \sum_{k=1}^{\infty} Q^{2k-1}(\widehat{g}_1) z^{2k-1}$$

where $P(z) \in \mathbb{Z}[x; p_g, g \in S]$ is an integer polynomial of x and p_g, $g \in S$.

Indeed, all the values $Q^k(g) \in \mathbb{Z}[p_g, g \in S]$ for any $k > 0$ and $g \in G$. Take $P(z)$ to be the difference between $s(z)$ and all the other series on the right hand side. We get the degree of $P(z)$ is at most n_0 (see above) which proves the above result.

Consider a set of formal commutative variables $x_{i,j}$, $1 \leq i, j \leq N$. For any set I, $I \subset [N]$ define a formal sum

$$R(i, j; I) = \sum_{k \geq 0, (i_1, \ldots, i_k) \in I^k} x_{i,i_1} \cdot x_{i_1,i_2} \cdot \ldots \cdot x_{i_k,j}$$

From the elementary combinatorial considerations (see e.g. [Stan], §4.7) we have:

$$R(i, j; I) = x_{i,j} + \sum_{l,m \in I} x_{i,l} \left(\frac{1}{1 - (x_{u,v})_{u,v \in I}} \right)_{l,m} x_{m,j}$$

where $M_{l,m}$ is a matrix element in column m and row l.

Therefore the series $R(i, j; I)$ can be presented as a ratio of two polynomials of $(x_{i,j})$:

$$R(i, j; I) = x_{i,j} + \sum_{l,m \in I} x_{i,l} \, \frac{(-1)^{l+m} \det(id - (x_{u,v})_{u \in I \setminus \{l\}, v \in I \setminus \{m\}})}{\det(id - (x_{u,v})_{u,v \in I})} \, x_{m,j}$$

(see e.g. [Stan], §4.7).

Let $x_{i,j} = b_{i,j}\, z^2$, where $B = (a_{i,j}) = A^2$ and A is an $N \times N$ - matrix as in the proof of the Theorem ?. Let $\widehat{g}_1 = g_i$, $\widehat{g}_2 = g_j$. We have

$$\sum_{k=1}^{\infty} Q^{2k}(\widehat{g}_1)\, z^{2\,k} = R(1, i; [N])$$

$$\sum_{k=1}^{\infty} Q^{2k-1}(\widehat{g}_2) = \sum_{l=1}^{N} a_{1,l}\, z\, R(l, j; [N])$$

This implies that each of the sums on the left hand side is a ratio of two polynomials in $\mathbb{Z}[x; p_g, g \in S]$. Combining with the above observations prove the result. $\square$

From Theorem 2.3.2 we immediately get a Theorem 2.2.3. Indeed, take the root of the denominator with the smallest absolute value ρ, which is also the radius of convergence of $s(x)$. Then $s_k \sim C\, \rho^{-k}$. $\square$

Example 2.3.3 Let $G = \mathbb{Z}_2$ as in the Example 2.2.12. Denote $g_1 = e$, $g_2 = a$. We have

$$A = \begin{pmatrix} \frac{1}{2} + \frac{p}{2} & \frac{1}{2} - \frac{p}{2} \\ \frac{1}{2} - \frac{p}{2} & \frac{1}{2} + \frac{p}{2} \end{pmatrix}$$

Since g_2 is an extremal element, we

$$s(z) = \frac{1}{1-z} - 2 \sum_{k=1}^{\infty} Q^k(g_2) = \frac{1}{1-z} - 2\, R(1, 2; [2])$$

We have

$$R(1, 2; [2]) = \frac{(-1)^{1+2}(-a_{1,2})}{\det(id - z\, A)}$$

$$\det(id - z\, A) = \left(1 - z\, \left(\frac{1}{2} + \frac{p}{2}\right)\right)^2 - \left(z\, \left(\frac{1}{2} - \frac{p}{2}\right)\right)^2 = (1-z)(1-z\,p)$$

and finally

$$s(z) = \frac{1}{1-z} - 2\, \frac{z\, \left(\frac{1}{2} + \frac{p}{2}\right)}{(1-z)\,(1-z\,p)} = \frac{(1-z\,p) - z\,(1-p)}{(1-z)\,(1-z\,p)} = \frac{1}{1-z\,p}$$

which agrees with our computations in Example 2.2.12.

Definition 2.3.4 The random walk $\mathcal{W} = (G, S, \mathbf{P})$ is called *rational* if $\mathbf{p}(s) \in \mathbb{Q}$ for all $s \in S$.

Corollary 2.3.5 Let $\mathcal{W}$ be a rational directed random walk. Then the separation series $s(x)$ is a ratio of two polynomials with rational coefficients.

Proof Obvious. $\square$

Corollary 2.3.6 Let $s = s(1)$ be the total separation of the rational directed random walk. Then s is a rational number.

Proof By the Lemma 2.2.5 the separation series $s(z)$ converges at 1 and therefore defines an analytic function on a disk $\{|z| < 1\}$. By Abel's Theorem (see e.g. [WW], §3.71 we have

$$\lim_{x \to 1-} s(x) = s(1) = s$$

From here and the Corollary 2.3.5 we get $s \in \mathbb{Q}$. $\square$

Remark 2.3.7 Asymptotic orders on groups were studied by various people (see [D] for references). When it is preserved from the beginning it is called *monotonicity* (see [D], §3*C*). For a random walk on the symmetric group generated by all transpositions it was computed by Lulov (see [Lulov], §8.2). In this case the asymptotic order becomes a reverse lexicographic on partitions which enumerate conjugacy classes. In this case the hitting time and the average hitting time was computed in [FOW].

The matrix X is usually called the *transfer matrix* (see e.g. [Stan], §4.7). Functions $R(i,j;I)$ are sometimes called *graph functions* and have numerous applications in Combinatorics and Algebra (see e.g. [Stan, Lall]). They have a noncommutative generalization which give analogs of the above results for a general class of Markov chains (see [GR, PPR]).

2.4 Other mixing times.

Definition 2.4.1 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk, Q^k be a probability distribution after $k > 0$ steps. Define the *total variation distance* tv_k as

$$tv_k = \frac{1}{2} \sum_{g \in G} \left| Q^k(g) - \frac{1}{|G|} \right|$$

The total variation distance is the most commonly used in the Markov chains setting. See [D], §3*B*, [AF], §2.6 for other definitions, generalizations, known results and references concerning the total variation distance.

Theorem 2.4.2 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk. Then for all $k \geq 0$

1) $tv_k = |G|\, Q^{2k}(e) - 1$

2) $1 - \sqrt{1 - s_{2k}} \leq tv_k \leq s_k, \quad k \geq n_{\frac{1}{2}}$

The second part of Theorem 2.4.2 is due to Aldous and Diaconis (see [AD2]). Proof of the Theorem in a more general setting can be found in [AD2], [D], §4*C*, [AF], §4.3. As a consequence of Theorem 2.4.2 we get bounds on the total variation distance from our bounds on separation distance.

Definition 2.4.3 The *total variation series* $tv(x)$ and the *average separation series* $as(x)$ can be defined as follows:

$$vt(x) = 1 + \sum_{k=1}^{\infty} \left(|G|\, Q^{2\,k}(e) - 1 \right) x^k$$

$$as(x) = 1 + \frac{1}{|G|} \sum_{g \in G} \sum_{k=1}^{\infty} \left| 1 - |G|\, Q^k(g) \right| x^k$$

Denote $vt = vt(1)$, $as = as(1)$.

Definition 2.4.4 Call the *hitting time* (*ht*) the maximum over all $g \in G$ of the mean time for a walk to *hit* g i.e. to get to g for the first time. The *average hitting time* (*aht*) is the mean time for a walk to hit a randomly chosen state. The *cover time* (*ct*) is the mean time to hit all the elements.

See [AF], §4, 6 for the definitions, generalizations, numerous results and examples of computation of the above "times". Later on, in §3.6 we will find some inequalities between them and the total separation.

Theorem 2.4.5 Let $\mathcal{W}$ be a rational directed random walk. Then the numbers $as = as(1)$, $tv = tv(1)$, *ht*, *aht* and *ct* are all rational.

Proof The proofs of all five results are almost exactly the same as the for the total separation (see §2.3).

There exist an analog of Proposition 2.2.4 and Lemma 2.2.5 (see §2.2) for the total variation distance (see e.g. [D], §3*B* and references there). This gives us convergence as $z \to 1-$ and finiteness of $tv(1)$. Analogously for $as(z)$ we get convergence at 1 since by definition the separation distance gives always the minimum of the difference $(1 - |G|\, Q^k(g))$ among all $g \in G$ and the total variation gives the maximum (see a short proof of the latter result in [ASE], p. 139). Presenting both series $vt(z)$ and $as(z)$ in terms of R we get the result.

In case of *ht*, *aht* and *cv* we do not have a convergence problem. Let $R(i, j; I)(x)$ denotes a series in x introduced in the proof of the Theorem 2.3.2. Observe that $R(1, i; [N] \setminus \{i\})(x)$ is the generating function by the number for steps of the probability of the walk hitting i for the first time. By definition of expectation, we have

$$ht = \max_{i \in [n]} R'(1, i; [N] \setminus \{i\})(1)$$

$$aht = \frac{1}{|G|} \sum_{i \in [N]} R'(1, i; [N] \setminus \{i\})(1)$$

This proves the result in case of *ht* and *aht*.

As for the cover time, the expansion in terms of R is somewhat tricky. Define $ct(z)$ the generating function for probability of covering all the group elements after k steps. We have

$$ct(z) = \sum_{\sigma \in S_{N-1}} R(1, 1+\sigma(1); \{1\}) \cdot R(1+\sigma(1), 1+\sigma(2); \{1, 1+\sigma(1)\}) \cdot \ldots \\ \cdot R\left(1+\sigma(N-2), 1+\sigma(N-1); [N] \setminus \{1+\sigma(N-1)\}\right)$$

Indeed, we start at e. First we wait till we hit the second element, then the third, etc, until we hit each of the elements. Summing over all the permutations of $\{2, \ldots, N\}$ we get the above formula.

Now since $cv = cv'(1)$ we get $cv \in \mathbb{Q}$ $\square$

Example 2.4.6 Suppose $G = \mathbb{Z}_4$ and $\mathcal{W}$ is a random walk as in the Example 2.1.1. Let us compute *ht*, *aht* and *ct*.

Denote E_1, E_2 the mean time to hit 1 and 2 respectively. The expected number of steps needed for a walk to hit either 1 or -1 is 2. From there it takes an expected 2 steps to move to either 0 or 2 with equal probability. This gives us

$$E_2 = 2 + 2 + \frac{1}{2}\, E_2$$

$$E_2 = 2\,(2+2) = 8$$

Analogously, it takes on average 2 steps for a walk to move to either 1 or -1 with equal probability; and from -1 it takes E_2 steps to get to 1. Therefore

$$E_1 = 2 + \frac{1}{2}\, E_2 = 6$$

Therefore

$$ht = 8, \qquad aht = \frac{1}{4}\,(0+6+8+6) = 5$$

For the cover time, it takes on average 2 steps to cover two elements, 4 more steps to cover the third element (by the symmetry) and E_1 steps to get to that last element. This gives us

$$ct = 2 + 4 + 6 = 12$$

2.5 Examples of random walks on groups.

Example 2.5.1 Let $G = \mathbb{Z}_m$, $S = \{0, \pm 1\}$, $\mathbf{p}(0) = \frac{1}{2}$, $\mathbf{p}(\pm 1) = \frac{1}{4}$. Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the *standard* random walk on the cyclic group $\mathbb{Z}_m$. In §4.1 we show that the total separation $s \sim \frac{1}{6}\, m^2$.

This is probably the oldest and most thoroughly studied random walk on a finite group. See [DF] for the modern treatment of this walk and bounds on separation distance.

Example 2.5.2 Let $G = \mathbb{Z}_2^n$, $S = \{e, s_i = (0, \ldots, 1_i, \ldots, 0), 1 \le i \le n\}$, $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(s_i) = \frac{1}{2\,n}$. Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the *standard* random walk on the n-cube $\mathbb{Z}_2^n$. In §4.3 we show that the total separation $s \sim n \ln(n)$.

This is another "old" example of a random walk. Detailed analysis of this walk can be found in [DGM].

Example 2.5.3 Let $G = S_n$. Take all the transpositions (i, j) to be elements of a set of generators S with uniform probability distribution $\mathbf{P}$ defined as follows:

$$\mathbf{p}(e) = \frac{1}{2}, \;\; \mathbf{p}(i,j) = \frac{1}{n\,(n-1)}$$

Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the random walk on S_n generated by all transpositions. In §5.2 we show that the total separation $s = O(n\, log(n))$.

This example was introduced and studied by Diaconis and Shahshahani in [DSh1], where they used a technique of bounding the total variation distance by considering the characters of the symmetric group. A simplified version of the proof can be found in [D], §3D.

Example 2.5.4 Let $G = S_n$. Take S to be a set of Coxeter generators i.e. transpositions $(i, i+1)$ (see §1.2). Let $\mathbf{P}$ be a uniform probability distribution:

$$\mathbf{p}(e) = \frac{1}{2},\ \mathbf{p}(i, i+1) = \frac{1}{2\,(n-1)}$$

for all $1 \leq i \leq n-1$. Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the random walk on S_n generated by adjacent transpositions. It is known (see [A, DSC]) that the total separation $s = O(n^3\, log(n))$. In §5.3 we show that $s = O(n^4)$ by a probabilistic argument.

This random walk was introduced and analyzed by Aldous in [A1] using a coupling argument. Later in [DSC] Diaconis and Saloff-Coste used a comparison technique to get similar bounds.

Example 2.5.5 Let $G = S_n$. Take S to be a set of *star* transpositions i.e. $S = \{e, (1, i); i = 2, \ldots, n\}$. Consider a uniform probability distribution $\mathbf{P}$ defined as

$$\mathbf{p}(e) = \frac{1}{2},\quad \mathbf{p}(1, i) = \frac{1}{2\,(n-1)}$$

for all $2 \leq i \leq n$. Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the random walk on S_n generated by star transpositions. In §5.1 we show that the total separation $s = O(n\, log(n))$.

This walk was studied by Flatto, Odlyzko and Wales in [FOW]. Comparison arguments in [DSC] give another proof of the result.

Example 2.5.6 Let $G = B_n$. Take S to be a set of *star* transpositions i.e. $S = \{e, \theta_1, \theta_i, (1, i); i = 2, \ldots, n\}$. Consider a uniform probability distribution $\mathbf{P}$ defined as

$$\mathbf{p}(e) = \frac{1}{2},\ \mathbf{p}(\theta_1) = \mathbf{p}(\theta_i) = \frac{1}{4\,n},\ \mathbf{p}(1, i) = \frac{1}{4\,(n-1)}$$

for all $2 \leq i \leq n$. Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the random walk on B_n generated by star transpositions. In §4.5 we show that the total separation $s = O(n\, log(n))$.

Example 2.5.7 Let $G = B_n$. Take S to be a set of the adjacent transpositions i.e. $S = \{e, \theta_n, \theta_i, (i, i+1); i = 1, \ldots, n-1\}$. Consider a uniform probability distribution $\mathbf{P}$ defined as

$$\mathbf{p}(e) = \frac{1}{2},\ \mathbf{p}(\theta_n) = \mathbf{p}(\theta_i) = \frac{1}{4\,n},\ \mathbf{p}(1, i) = \frac{1}{4\,(n-1)}$$

for all $1 \leq i \leq n-1$. Call the walk $\mathcal{W} = (G, S, \mathbf{P})$ the random walk on B_n generated by adjacent transpositions. In §4.5 we show that the total separation $s = O(n^4)$.

Example 2.5.8 Let $G = U(n, \mathbb{F}_q)$. For any $1 \leq i < j \leq n$, $a \in \mathbb{F}_q$ define an upper triangular matrix $R(i, j; a)$ with ones on the diagonal, a in place (i, j) and zeros elsewhere. Let

$$S = \{R(i, j; a),\ 1 \leq i < j \leq n,\, a \in \mathbb{F}_q\}$$

be a symmetric set of generators. Define a probability distribution $\mathbf{P}$ as

$$\mathbf{p}(e) = \frac{1}{q},\quad \mathbf{p}(R(i, j; a)) = \frac{1}{q\,\binom{n}{2}},\ a \neq 0$$

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$. We shall call this random walk the standard random walk on the upper triangular matrices. In §6.3 we prove that the total separation $s = O(n\,log(n))$.

Example 2.5.9 Let $G = U(n, \mathbb{F}_q)$, $S = \{e, R(i,j;\pm 1),\ 1 \le i < j \le n\}$. Define the probability distribution $\mathbf{P}$ as

$$\mathbf{p}(e) = \frac{1}{2}, \quad \mathbf{p}(R(i,j;\pm 1)) = \frac{1}{2\,n\,(n-1)},\ 1 \le i < j \le n$$

Consider a lazy random walk $\mathcal{W} = (G, S, \mathbf{P})$. In §6.3 we obtain the following bound on the total separation: $s = O(n^2\,log(n))$

2.6 General bounds on the separation distance.

Let G be a finite group, S - its symmetric set of generators and $d = d_S$ its diameter. For any k we have

$$s \le k \cdot (1 + s_k + s_k^2 + \dots) = \frac{k}{1 - s_k}$$

Denote $\widehat{g}_k$ a k-minimal element, $m_k = Q^k(\widehat{g}_k)$. In other words,

$$m_k = \min_{g \in G} Q^k(g)$$

If $k \ge d_S$ we have

$$0 < m_k \le \frac{1}{|G|}$$

and

$$s_k = 1 - |G| \cdot Q,\ 0 \le s_k < 1$$

Substituting here the value of s_k we get the following result:

Theorem 2.6.1 For any directed random walk $\mathcal{W}$ and $k \ge d_S$ we have

$$s \le \frac{k}{|G|\,Q^k(\widehat{g}_k)}$$

where $\widehat{g}_k$ is a k-minimal element.

In particular, when $k = d_S$ we get

$$s \le \frac{d_S}{|G|\,Q^{d_S}(\widehat{g}_{d_S})}$$

Note that $Q^k(\widehat{g}_k) \ge \frac{c}{|G|}$ implies $s_n \le (1-c)^{\lfloor n/k \rfloor}$. This is called Doeblins condition. Later on we are going to generalize this result for various stopping times on groups (see §3.5).

Example 2.6.2 Let $G \simeq \mathbb{Z}_2$, $S = G = \{e, a\}$, $\mathbf{p}(a) = p/2$, $\mathbf{p}(e) = 1 - p/2$, $p < 1$. We have $Q = p/2$, $d_S = 1$. The theorem therefore gives us an upper bound $s \leq \frac{1}{p}$ which is tight since in Example 2.2.12 we showed that $s = \frac{1}{p}$.

Example 2.6.3 let $G = \mathbb{Z}_4$ and a walk as in Example 2.1.1. We have $d_S = 2$, $D_S = \{2\}$, $Q = Q^2(2) = 1/8$. Therefore the total separation is bounded by

$$s \leq \frac{2}{4\,\frac{1}{8}} = 4$$

Recall that $s = 3$ is the right value of the total separation (see Example 2.2.12).

Example 2.6.4 Let G be a finite group, $N = |G| - 1$, and $\mathcal{W} = (G, S, \mathbf{P})$ a random walk as in Example 2.2.14. We have $d_S = 2$, $D_S = \{a\}$, $Q = Q^2(a) = \frac{N-1}{N^2}$. Therefore the total separation is bounded by

$$s \leq \frac{2}{(N+1)\cdot\frac{N-1}{N^2}} = \frac{2\,N^2}{N^2-1}$$

which is a tight upper bound (see Example 2.2.14).

Example 2.6.5 Let $G = U(3; \mathbb{F}_p)$, p is a prime, the group of upper triangular matrices over the finite field with p elements. Consider the following set of generators S:

$$R_1(a) = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad R_2(b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

where $a, b \in \mathbb{F}_p$. Note that $R_1(e) = R_2(e)$ and $R_i(a)\,R_i(b) = R_i(a+b)$, $i = 1, 2$.

We claim that $d_S = 4$ and D_S contains the matrix

$$T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Indeed, T cannot be obtained as a product of any three matrices $R \in S$. Therefore $d_S > 3$. On the other hand we claim that the product $R_2(a)\,R_1(b)\,R_2(c)\,R_1(d)$ contains each matrix at least $p-1$ times. Solve the equations

$$R_2(a)\,R_1(b)\,R_2(c)\,R_1(d) = \begin{pmatrix} 1 & b+d & b\,c \\ 0 & 1 & a+c \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

We get $2\,p-1$ solutions if $z = 0$ and $p-1$ solutions otherwise. Analogously, for the product $R_1(a)\,R_2(b)\,R_1(c)\,R_2(d)$ we get $2\,p-1$ solutions if $z = x\,y$ and $p-1$ solutions otherwise. These are the only ways one could get a matrix T since any repetition is a sequence of indices of R's means that the product can be written as a product of only three such matrices.

Define a probability distribution $\mathbf{P}$ as follows:

$$\mathbf{p}(e) = \frac{1}{p}\,, \quad \mathbf{p}(R_1(a)) = \mathbf{p}(R_2(b)) = \frac{1}{2\,p},\ a, b \neq 0$$

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$. Each walk step can be described as follows:

- Flip a fair coin. Choose a random number $a \in \mathbb{F}_p$.
- If heads, apply generator $R_1(a)$
- If tails, apply generator $R_2(a)$

Lemma 2.6.6 For any $U \in U(3, \mathbb{F}_p)$

$$Q^4(U) \geq Q^4(T) = \frac{p-1}{8\,p^4}$$

This implies that T is a 4-minimal element. From here and Theorem 2.6.1 we immediately get a bound on the total separation for the random walk $\mathcal{W}$:

$$s \leq \frac{4}{p^3\,\frac{p-1}{8\,p^4}} = 32\,\frac{p}{p-1}$$

In particular, we get $s \leq 40$ when $p \geq 5$. If $p = 2$ we get $s \leq 64$ while the right value of the total separation in this case is $s = 11$ (see §4.1).

Proof of Lemma 2.6.6 We know that the matrix T occurs exactly $2\,(p-1)$ times as a product of four matrices $R_i(a)$, $i = 1, 2$, $a \in \mathbb{F}_p$. There are $(2\,p)^4$ possible products which gives us a value $Q^4(T) = \frac{p-1}{8\,p^4}$.

Observe now that $2\,(p-1)$ times is the minimal number each matrix U occurs in such a product. This proves the inequality. $\square$

Example 2.6.7 Let $G = S_n$. Consider the set S of all permutations without fixed points. Define a uniform random walk $\mathcal{W} = (G, S, \mathbf{P})$ with $\mathbf{p}(e) = 0$, $\mathbf{p}(s) = \frac{1}{|S|}$, $s \in S$. By use of the inclusion-exclusion principle one can show that

$$|S| = \left\lfloor \frac{n! + 1}{e} \right\rfloor$$

(see e.g. [Stan], §2.2.1). Also $d_S = 2$ for all $n \geq 4$, so one can expect random walk $\mathcal{W}$ to be rapidly mixing. Here is a result we can use to prove rapid mixing.

Lemma 2.6.8 Let $\sigma = (2, 3, \ldots, n, 1) \in S_n$ - n-cycle, $n \geq 4$. Then

1) σ is a 2-minimal element for a random walk $\mathcal{W}$

2) $Q^2(\sigma) \geq \frac{1}{n!}\left(1 - \frac{1}{n-1}\right)$, $n \geq 5$.

From Lemma 2.6.8 and Theorem 2.6.1 we get a bound on the total separation for this random walk on S_n, $n \geq 5$

$$s \leq \frac{2}{n! \cdot \frac{1}{n!}\left(1 - \frac{1}{n-1}\right)} = 2 + \frac{2}{n-2}$$

Lemma 2.6.8 is proved in [P1] by a careful analysis of certain rook placements.

When dealing with separation distance it is often hard to even locate the k-minimal element $\widehat{g}_k$ and also hard to compute the value $Q^k(\widehat{g}_k)$. Without a lower bound on $Q = Q^{d_S}(\widehat{g}_{d_S})$ Theorem 2.6.1 is useless. In some cases, however, the following conjecture helps to determine the d_S-minimal elements.

Conjecture 2.6.9 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk and $\mathbf{p}(e) \geq \frac{1}{2}$. Then the diameter set D_S contains a d_S-minimal element $\widehat{g}_{d_S}$.

We do not attempt to prove this conjecture, although we have every reason to believe that it is true. Assuming that it is true, we can get some exponential bounds on the total separation (see two examples below)

Example 2.6.10 Let $G = \mathbb{Z}_{2\,m}$. Let $\mathcal{W} = (G, S, \mathbf{P})$ be standard random walk on $\mathbb{Z}_{2\,m}$ defined as in Example 2.5.1. We have $d_S = m$, $D_S = \{m\}$, $Q = Q^m(m) = 2\,\frac{1}{4^m} = 2^{1-2\,m}$. Therefore assuming that Conjecture 2.6.9 holds, the total separation is bounded as

$$s \leq \frac{m}{2\,m\,2^{1-2\,m}} = 4^{m-1}$$

Recall (see Example 2.5.1) that the real bound is asymptotically much smaller:

$$s = O(m^2)$$

Example 2.6.11 Let $G = \mathbb{Z}_2^n$, Let $\mathcal{W} = (G, S, \mathbf{P})$ be standard random walk on $\mathbb{Z}_2^n$ defined as in Example 2.5.2. Then $d_S = n$, $D_S = \{(1, \ldots, 1)\}$, $Q = Q^n(1, \ldots, 1)$.

Observe that the number of minimal paths of length n going from $e = (0, \ldots, 0)$ to $\widehat{g} = (1, \ldots, 1)$ in a Cayley graph Γ is exactly $r_n = n!$. Indeed, we need to take one step in direction of each coordinate which makes the number of steps equal to the number of coordinates permutations or $n!$. From part *i)* of Theorem 2.1.5 we get

$$Q = Q^m(1, \ldots, 1) = \frac{n!}{(2\,n)^n}$$

It gives us a bound for the total variation distance

$$s \leq \frac{n}{2^n\,n!\,2^{-n}\,n^{-n}} = \frac{n^{n+1}}{n!}$$

Recall the Stirling formula

$$n! \sim \sqrt{2\,\pi\,n}\,\left(\frac{n}{e}\right)^n$$

Assuming that Conjecture 2.6.9 holds, we have

$$s \leq C \sqrt{n}\, e^n$$

which is a bad bound compared to the right bound

$$s \sim 2\, n\, log(n)$$

.

2.7 The Diameter Conjecture.

There seems to be a very strong connection between the diameter d_S and the total separation. We would like to propose the following conjecture.

Diameter Conjecture 2.7.1
There exist universal constants $C, \varepsilon > 0$ such that for any finite group G and a directed random walk $\mathcal{W} = (G, S, \mathbf{P})$ we have

$$s < C\, \frac{d_S^{\varepsilon}}{p}$$

where $p = \min_{s \in S} \mathbf{p}(s)$.

We actually believe that the constant ε is rather small, i.e. $\varepsilon < 8$. There are few partial results known (see [DSC], [Chung]). The real power of the above conjecture can be seen in case of a simple group G when it can be combined with the following conjecture of Babai (see [Babai]).

Conjecture 2.7.2 (Babai)
There exist universal constants $C, \epsilon > 0$ such that for any nonabelian simple group G and its set of generators S we have:

$$d_S < C\, (\log(|G|))^{\epsilon}$$

Unfortunately this conjecture is open even for $G = A_n$, as $n \to \infty$ (see [Babai]). Recall that there exist a classification of the simple groups. This gives us a hope that the Conjecture 2.7.2 will be proven eventually.

There is an extensive literature on the bounds for the diameter d_S of a permutation groups $G \subset S_n$ (see references in [Babai, DM]). We would like to recall the following result.

Theorem 2.7.3 There exist universal constants C, C', C_k, $k > 0$ such that for any permutation group $G \subset S_n$ and its set of generators S
1) $d_S < C_k\, n^2$, if S consists of cycles of length $\leq k$.
2) $d_S < C\, n^{2\,k}$, if S consists of elements of degree $\leq k$.
3) $d_S < C'\, e^{\sqrt{n\ ln(n)}(1+o(1))}$, if no assumptions on S is made.

The first part of this Theorem is due to Driscoll and Furst, the second part is due to McKenzie and the last part - to Babai and Seress (see [Babai] for the references).

For an abstract group G denote $n(G)$ the minimal n for which G is a subgroup of S_n. Of course, $n \leq |G|$, but in our examples (see below) we will show that $n(G)$ is much smaller. Assuming Diameter Conjecture, the Theorem 2.7.3 basically gives us polynomial bounds in most of the interesting cases.

Example 2.7.4 As in Example 2.2.12, consider a group $G \simeq \mathbb{Z}_2$ with $S = G = \{e, a\}$, $\mathbf{p}(a) = \frac{p}{2}$, $\mathbf{p}(e) = 1 - \frac{p}{2}$, $p < 1$. For this random walk we get $s = \frac{1}{p}$ (see Example 2.2.12). This proves that we cannot get rid of p in the statement of the Diameter Conjecture.

Example 2.7.5 Let $G = \mathbb{Z}_m$, $\mathcal{W} = (G, S, \mathbf{P})$ be a standard random walk on $\mathbb{Z}_m$ (see Example 2.5.1. We have $d_S = \lfloor \frac{m}{2} \rfloor$, $p = 1/4$. The Diameter Conjecture gives us a polynomial bound

$$s < C\, m^{c'}$$

while the right bound is

$$s = O(m^2)$$

Consider the first k primes $p_1, p_2, \ldots, p_k$ and take $m = p_1 \cdot \ldots \cdot p_k$, $n = p_1 + \cdots + p_k$. A permutation $\sigma \in S_n$ with cycles of lengths p_1, p_2, ... , p_k generates a cyclic group $\mathbb{Z}_m$. By use of the standard asymptotic technique and the Prime Number Theorem (see [Odl], Ex 5.1) one can show that the bound in part *3)* of the Theorem 2.7.3 is tight (see [Babai, Babai-Seress]).

Example 2.7.6 Let $G = \mathbb{Z}_2^m$, $\mathcal{W} = (G, S, \mathbf{P})$ as in Example 2.5.2. We have $n(G) \leq 2\, m$. Indeed, $\mathbb{Z}_2^m$ is isomorphic to a subgroup $H \subset S_{2\, m}$ generated by transpositions $(2\, i, 2\, i + 1)$, $i = 1, \ldots, m$. In our example $p = \frac{1}{2\, m}$, $d_S = m$ and the Diameter Conjecture gives us again a polynomial bound

$$s < C\, m^{c'}$$

when the right bound is

$$s \approx C\, m\, log(m))$$

Note that in general,

$$n(G_1 \times G_2) \leq n(G_1) + n(G_2)$$

and if G_1, G_2 are generated by cycles of length $\leq k$, so is $G_1 \times G_2$. We will come back to this example in §4.3.

Example 2.7.7 Let G be the symmetric group S_n. We consider many different sets of generators (see examples $2.5.3 - 5$). Suppose S_n is generated by a set S of all transpositions. $|S| = O(n^2)$. The Diameter Conjecture gives us a polynomial bound

$$s < C\, n^{c'}$$

On the other hand, since A_n is simple and normal in S_n one can use Conjecture 2.7.2 to get a polynomial bound. Indeed, we have a normal tower $S_n \supset A_n \supset \{e\}$. By Stirling's formula $log(|A_n|) = O(n\,log(n))$. Since the factor $S_n/A_n \simeq \mathbb{Z}_2$, we get a polynomial bound for the total separation.

Observe that we can use the Diameter Conjecture combined with Conjecture 2.7.2 for *any* set of generators. In case $|S| = O(1)$ there is a conjecture of Persi Diaconis (see [DP]) which can be formulated in term of the total separation as follows.

Conjecture 2.7.8 (Diaconis)
Fix $k > 1$. Suppose for all $n > n_0$ there is a set of k generators $S^n = \{s_1^n, \ldots, s_k^n\}$ of the symmetric group S_n. Define $\mathcal{W}_n = (S_n, S^n, \mathbf{P})$, where $\mathbf{p}(s_k^n) = 1/k$. Then there is a constant $C > 0$ such that

$$s < C\,n^3\,log(n)$$

for all $n > n_0$, where s is a total separation for a random walk $\mathcal{W}_n$.

In [DSC] Diaconis and Saloff-Coste show that this is the right bound in case of a generating set consisting of just four generators: an identity, a transposition, an n-cycle and its inverse.

Example 2.7.9 Let $G = B_n$. It is easy to see that $n(B_n) \subset S_{2\,n}$. Indeed, simply take all the permutations $\sigma \in S_{2\,n}$ such that $\sigma(j) = j (\mathrm{mod}\ n)$, $1 \le j \le 2\,n$. In general, for any wreath product $G \ltimes H^n$ we can use analogous argument to prove that

$$n(G \times H) \le n \cdot n(H)$$

Let $\mathcal{W}_1$, $\mathcal{W}_2$ be the random walks on B_n generated by the star transpositions and by the adjacent transpositions. Since all the generators are transpositions, we can use part *1)* of the Theorem 2.7.3 and the Diameter Conjecture. This gives us a polynomial bound

$$s < C\,n^{c'}$$

Example 2.7.10 Let $G = GL(n; \mathbb{F}_q)$, $n \ge 3$. There are great many different sets of generators of this group, some of them consisting of as little as two elements (see [CM], §7.2). One can consider a uniform random walk generated by a set of $O(1)$ elements as either q or n tends to infinity.

Since $PSL(n; \mathbb{F}_q)$ is a simple group and $\frac{|GL(n;\mathbb{F}_q)|}{|PSL(n;\mathbb{F}_q)|} \le q^2$ one can reduce the problem of estimating the diameter on $GL(n; \mathbb{F}_q)$ to the analogous problem for $PSL(n; \mathbb{F}_q)$. By Conjecture 2.7.2 we get:

$$d_S < C\,log^{\epsilon}(|PSL(n, \mathbb{F}_q)|) = C\,n^{2\,\epsilon}\,log^{\epsilon}(q)$$

for any set S of generators of a group $PSL(n; \mathbb{F}_q)$. Now we can apply the Diameter Conjecture to get a bound on the total separation s.

Among very few known results regarding random walks on linear groups note a result by Hilderbrand (see [Hild]). He studied a uniform random walk on $SL(n, \mathbb{F}_q)$ with transvections $R(i, j; a)$, 1 $lei, j \le n$, $i \ne j$, $a \in \mathbb{F}_q$ as a generating set. It turns

out that we need $O(n)$ steps for this walk to get in random (see [Hild], [Gluck]). One can also apply a Diameter Conjecture in this case since by result of Jason Fulman the diameter in this case is at most $(3\,n-3)$ (see [Ful]).

In this work we do not further consider random walks on $GL(n;\mathbb{F}_q)$.

Example 2.7.11 Let $G = U(n;\mathbb{F}_p)$, p-prime, $S = \{e, R(i,i+1;\pm 1), 1 \leq i \leq n-1\}$. Define the probability distribution $\mathbf{P}$ as

$$\mathbf{p}(e) = \mathbf{p}(R(i,i+1;\pm 1)) = \frac{1}{2\,n-1)},\ 1 \leq i \leq n-1$$

Ellenberg in [Ell] finds the sharps bounds on the diameter d_S:

$$d_S = O(n\,p + n^2\,log(p))$$

as $n+p \to \infty$. One can use the Diameter Conjecture in order to get a polynomial bound for a uniform random walk $\mathcal{W} = (G, S, \mathbf{P})$ on $U(n,\mathbb{F}_p)$.

The random walk $\mathcal{W} = (G, S, \mathbf{P})$ was studied by Diaconis and Saloff-Coste in [D-S-C] and later by Stong in [Stong1]. Stong shows that

$$s = O(p^2\,n^3\ \log(p))$$

as $n+p$ tends to infinity. We do not further consider this random walk in this work. Note, however, that the sharp bounds can be obtained by comparison this walk and the standard random walk on the upper triangular matrices.

3. Stopping times

3.1 Randomized stopping times.

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$. Define $\mathcal{X}$ to be the set of all possible walk paths

$$\mathcal{X} = \{X = (X_0, X_1, X_2, \dots)\}$$

Definition 3.1.1 A *stopping time* τ is a function

$$\tau : \mathcal{X} \to \mathbb{N} \cup \{\infty\}$$

which satisfies the following condition

- $\tau(X)$ does not depend on $X_{\tau(X)+1}$, $X_{\tau(X)+2}$, $\dots$

Here is one way to think about stopping times. Consider an algorithm which observes movements of a walk and according to some *stopping rule* stops the walk at one point. This stopping rule could be a complicated function of the states the walk passed, but it can't possibly depend on the states walk is going to pass after stopping since *they haven't been observed yet*.

Sometimes it is also useful to have the stopping rule depend not only on the states passed but also on some random events. Here is how it can be done.

Definition 3.1.2 A *randomized stopping time* τ is a function $\tau(X) = P$ where $X \in \mathcal{X}$, $P = P_X$ is a function $P_X : \mathbb{N} \cup \{\infty\} \to \mathbb{R}$, $P_X(i) = p_{X,i}$, satisfying the following conditions:

1) $p_{X,1} + p_{X,2} + \dots = 1$, $p_{X,i} \geq 0$, $i \in \mathbb{N} \cup \{\infty\}$, $X \in \mathcal{X}$

2) $p_{X,i} = p_{X',i}$ if $X_j = X'_j$ for all $1 \leq j \leq i$

In other words, for each walk path $X \in \mathcal{X}$ we now have a probability distribution P_X which says that after i steps we should stop with probability $p_{X,i}$. When a probability distribution P_X is concentrated at one point $i = \tau(X)$, i.e. $p_{X,i} = 1$ and $p_{X,j} = 0$ if $i \neq j$, we get a usual stopping time.

As before it is easier to think of a randomized stopping times in terms of an algorithm which observes movements of a walk and stops according to some *randomized stopping rule*. The stopping rule for a random walk $\mathcal{W} = (G, S, \mathbf{P})$ is designed as follows.

Algorithm G Suppose the walk X_t walked along the word $\omega_n \in S_n^{st}$. Flip an unfair coin with probability of heads $\mathfrak{P} = \mathfrak{P}(\omega)$.

- if heads, stop.
- if tails, walk one step (use Procedure **W**);
- Return to the beginning of the Algorithm.

Obviously the probability of heads $\mathfrak{P} = \mathfrak{P}(\omega)$ depends only on the states passed and the previous coin tossings. Therefore Algorithm **G** always defines a randomized stopping time.

Vice versa, each randomized stopping time has this algorithmic interpretation. To see that observe that probability of heads $\mathfrak{P}(\omega)$ in the Algorithm **G** is equal to the probability of stopping after the random walk X_t walked along the word ω condition on not stopping prior to that. We have

$$\mathfrak{P}(\omega_i) = \frac{p_{X,i}}{1 - p_{X,1} - p_{X,2} - \cdots - p_{X,i-1}}$$

By the argument above probabilities $\mathfrak{P}$ define a randomized stopping time. Checking that it is the stopping time we started with is trivial.

Call a *stopping state* $\varrho(X) = X_{\tau(X)} \in G$ the state where the algorithm stops. We can think of $\tau(X)$ and $\varrho(X)$ as random variables defined on some probability space Ω which includes set of walk path $\mathcal{X}$ and independent of $\mathcal{X}$ set of coin flipping outcomes.

We can think of the probabilities $\mathfrak{P}(\omega)$ as the values of a function $\mathfrak{P} : S^\star \to [0, 1]$. We call $\mathfrak{P}$ a *stopping probability function*. Denote $\mathfrak{Q}(\omega) = 1 - \mathfrak{P}(\omega)$ the probability of tails in Algorithm **G**.

Note that now we can define a randomized stopping time for random walks which start at a some state $g_0 \in G$, not necessarily at 0. Simply use Algorithm G with the same stopping probability function $\mathfrak{P}$.

Unless it leads to confusion, just for convenience we drop the "randomized" part in the "randomized stopping time" and "randomized stopping rule".

In the future we will always describe stopping times in term of certain algorithms which represent the stopping rules. We want to be able to construct new stopping rules out of those already known. In order to do that with each algorithm A we associate a variable **A** which takes value 1 if the algorithm stops the walk at this time or before, and 0 otherwise. In other words our stopping time can be presented as an algorithm:

Algorithm A
- Walk till $\mathbf{A} = 1$.
- Stop.

We call a variable **A** a *stopping call*.

Now we can write

$$\tau_1 = \tau + 1$$

This means that the new rule is defined as follows

- walk one more step than in the rule τ

In terms of the algorithm this rule can be formalized as follows.

Algorithm A′
- Walk till $\mathbf{A} = 1$.
- Walk one more step. Stop.

We can also define $\tau_1 + \tau_2$, $min(\tau_1, \tau_2)$, $max(\tau_1, \tau_2)$ by the following rules:

- Walk till $\mathbf{A}_1 = 1$. Then walk till $\mathbf{A}_2 = 1$. Stop.
- Walk till either $\mathbf{A}_1 = 1$ or $\mathbf{A}_2 = 1$. Stop.
- Walk till both $\mathbf{A}_1 = 1$ and $\mathbf{A}_2 = 1$. Stop.

3.2 Strong uniform times.

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$ and a stopping time τ.

Definition 3.2.1 Stopping time τ is called *uniform* if for all $g \in G$

$$P(\varrho = g) = \frac{1}{|G|}$$

where ϱ is the stopping state.

In other words we want our stopping state ϱ to be uniformly distributed on a group G. In order to able to compute the separation distance of a walk we actually need to impose a stronger condition.

Definition 3.2.2 A stopping time τ is called *strong uniform* if for all $g \in G$ and

$$P(\varrho = g | \tau = k) = \frac{1}{|G|}$$

Note that if τ is strong uniform, so is $\tau + 1$. Not all uniform times are strong uniform. Here is an example.

Example 3.2.3 Consider a standard random walk on $\mathbb{Z}_4$ (see Example 2.1.1). Start at $e = 0$. Recall that at each time we either stay with probability $\frac{1}{2}$ or move to the nearest number with equal probability. Define the stopping time τ by the following algorithm.

Algorithm 3.2.4
Choose a random state and walk till we get there. Stop.

By definition it is a uniform stopping time. On the other hand is not a strong uniform time. Indeed, if we know the walk stopped after one step, it couldn't possibly be the 2, while both 1 and 3 can occur.

Example 3.2.5 Now we present the correct construction of the strong uniform time τ on $\mathbb{Z}_4$.

Algorithm 3.2.6

- Walk till you hit either 1 or 3.
- Do one more step and stop.

Here is how this stopping rule can be translated from "English" into our language:

- If the walk is at state $X_k = g$ after k steps let
  $\mathfrak{P} = 1$ if $X_{k-1} = \pm 1$
  $\mathfrak{P} = 0$ otherwise.

Let us prove that this stopping time is indeed strong uniform. Suppose we stopped after $\tau = k$ steps. By symmetry

$$P(X_{k-1} = 1 | \tau = k) = P(X_{k-1} = 3 | \tau = k) = \frac{1}{2}$$

Now for all $i = 0, \dots, 3$ we can compute the probabilities of stopping at $i \in \mathbb{Z}_4$. We have

$$P(X_t = i|\tau = k) = \frac{1}{2}P(X_{k-1} = i|\tau = k) + \frac{1}{4}P(X_{k-1} = i \pm 1|\tau = k) = \frac{1}{4},$$

$i = 0, \dots, 3$. Therefore τ is indeed strong uniform.

The following result is the main reason for study of the strong uniform times. It is proved in Proposition 3.3.2 and Remark 3.3.3 below.

Theorem 3.2.7 (Aldous, Diaconis)
Let $\mathcal{W} = (G, S, P)$ be a random walk on a finite group G. Then
1) for any strong uniform time τ and any $k \geq 0$ we have

$$s_k \leq P(\tau > k)$$

2) there exist a strong uniform time $\widetilde{\tau}$ such that for any $k \geq 0$ we have

$$s_k = P(\widetilde{\tau} > k)$$

Definition 3.2.8 A stopping time $\widetilde{\tau}$ is called *perfect* if it satisfies the equation in the second part of Theorem 3.2.7.

From Theorem 3.2.7 we immediately get

Corollary 3.2.9 Under the conditions of Theorem 3.2.7:
1) For any randomized strong uniform time τ we have

$$s \leq E(\tau)$$

2) There exist a perfect time $\widetilde{\tau}$. For this stopping time we have

$$s = E(\widetilde{\tau})$$

where $s = s(1)$ is a total separation (see §2.2).

Proof Indeed, by definition of the total separation we have

$$s = s_0 + s_1 + s_2 + \dots = P(\widetilde{\tau} > 0) + P(\widetilde{\tau} > 1) + P(\widetilde{\tau} > 2) + \dots$$
$$= 1 \cdot P(\widetilde{\tau} = 1) + 2 \cdot P(\widetilde{\tau} = 2) + 3 \cdot P(\widetilde{\tau} = 3) + \dots = E(\widetilde{\tau})$$

This finishes the proof of the second part of the Corollary. The proof of the first part is almost identical. □

Remark 3.2.10 Strong uniform times were introduced by Aldous and Diaconis in [AD1, AD2]. The results in this section can be also found in [D, AF].

Theorem 3.2.11 Let τ_1, τ_2 be randomized stopping times defined by their stopping probability functions $\mathfrak{P}_1$, $\mathfrak{P}_2$. If either τ_1 or τ_2 is strong uniform, then $\tau_1 + \tau_2$ is also strong uniform.

Proof Clear □

3.3 Properties of strong uniform times.

Proposition 3.3.1 For any strong uniform time τ, $g \in G$ and $k \geq 0$ we have

$$Q^k(g) = \frac{1}{|G|} P(\tau \leq k) + P(X_k = g, \tau > k)$$

Proof By definition we have

$$Q^k(g) = P(X_k = g) = \sum_{i=1}^{k} P(X_k = g, \tau = i) \ + P(X_k = g, \tau(X) > k)$$

For each of the terms in the last sum we have

$$\begin{aligned} P(X_k = g, \tau = i) &= \sum_{g' \in G} P(X_i = g' | \tau = i) \cdot P(\tau = i) \cdot P(X_k = g | X_i = g') \\ &= \frac{1}{|G|} P(\tau = i) \end{aligned}$$

Therefore the sum on the right hand side in the proposition is equal to

$$\sum_{i=1}^{k} \frac{1}{|G|} P(\tau = i) = \frac{1}{|G|} P(\tau \leq k)$$

This finishes the proof of the proposition. $\square$

Remark 3.3.2 Note that in the proof of the proposition we used the condition $P(\varrho = g | \tau = l) = \frac{1}{|G|}$ only for $1 \leq l \leq k$.

From Proposition 3.3.1 we immediately get the first part of Theorem 3.2.7. Indeed,

$$\begin{aligned} |G| \left(\frac{1}{|G|} - Q^k(g) \right) &= 1 - P(\tau \leq k) - |G| \cdot P(X_k = g, \tau > k) \\ &\leq P(\tau > k) \end{aligned}$$

Since separation is the maximum over all $g \in G$ we get

$$s_k = |G| \cdot \max_{g \in G} \left(\frac{1}{|G|} - Q^k(g) \right) \leq P(\tau > k)$$

which proves first part of Theorem 3.2.7.

In order to prove the second part of Theorem 3.2.7 observe that the equality holds if and only if for any $k > 0$ there exist an element $\widehat{g}_k$ such that $P(X_k = \widehat{g}_k, \tau > k) = 0$. In other words, if the walks gets to the state g after k steps, it stops there.

We can (in theory) determine these elements $\widehat{g}_k$. Since we do maximization over all group elements $g \in G$, we get

$$Q^k(\widehat{g}_k) \leq Q^k(g)$$

for any $g \in G$. Thus $\widehat{g}_k$ is actually a k-minimal element (see §2.1) and we obtained the following result.

Theorem 3.3.3
1) Let $(\widehat{g}_1, \widehat{g}_2, \dots)$ be a minimal sequence for a random walk $\mathcal{W} = (G, S, P)$. A strong uniform time τ is perfect if and only if the *stopping property* holds:

- whenever we have $X_k = \widehat{g}_k$, $k > 0$ we also have $\tau = k$.

2) If we have the stopping property for some strong uniform time τ and a sequence $(g_1, g_2, \dots)$ this implies that that τ is perfect and the sequence is minimal.

Example 3.3.4 Recall our example of a random walk on $\mathbb{Z}_4$. As we computed in Example 2.2.13 for each $k > 1$ there is exactly one k-minimal element $g_k = 2$. Since we already have a construction of a strong uniform time τ for this random walk (see Example 3.2.5), we can check whether this is a perfect time.

Since we start at 0, clearly the only way we could possibly get to 2 is by moving there from either 1 or 3. This means that once we get to 2 we always stop there (of course, we can also stop in other states). Therefore by the theorem above the sequence $(2, 2, 2, \dots)$ is indeed a minimal sequence and the strong uniform time τ is perfect.

From here and Theorem 3.2.7 we have

$$s_k = P(\tau > k) = P(X_1 = X_2 = \cdots = X_{k-1} = 0) = \frac{1}{2^{k-1}}$$

where $k > 0$. This agrees with our computations in Example 2.2.13.

Before we finish this section, we would like to give one more definition. As in the example above, in practice it sometimes happens that there is an element g which is k-minimal for all $k > 0$.

Definition 3.3.5 We say that a strong uniform time τ *respects* an element $\widehat{g} \in G$ if it always stops once it gets there.

By Theorem 3.3.3 τ is perfect and $\widehat{g}$ is extremal.

If it exists, an extremal element $\widehat{g}$ is usually easy to find. It must always belong to the diameter set D_S. Indeed, since $\widehat{g}$ is k-minimal for all $k \in \mathbb{N}$, it must be also $d_S - 1$-minimal, where d_S is a diameter (see §1.1). But $Q^{d_S-1}(g) > 0$ for all $g \in G \setminus D_S$ and $Q^{d_S-1}(g) = 0$ for all $g \in D_S$. Therefore we just proved

Proposition 3.3.6 Let $\widehat{g}$ be an extremal element of G. Then

$$\widehat{g} \in D_S$$

Unfortunately finding a suspect doesn't always help. Virtually all the random walks we are going to study are guileless although proving it is very difficult sometimes. In some cases the only existing proof is given by the construction of the perfect time which respects some specific extremal element.

3.4 Making a perfect time.

Now we present a construction of a stopping probability function $\mathfrak{P}$ which defines a perfect time for any directed random walk $\mathcal{W} = (G, S, \mathbf{P})$.

Definition 3.4.1 For any $g \in G$, $\omega \in S_k^\star(g)$ let

$$\mathfrak{P}(\omega) = \mathfrak{P}(g) = \frac{Q^k(\widehat{g}_k) - Q^{k-1}(\widehat{g}_{k-1})}{Q^k(g) - Q^{k-1}(\widehat{g}_{k-1})}$$

where $\widehat{g}_l$ is an l-minimal element, $l \geq 1$.

Theorem 3.4.2 The stopping probability function $\mathfrak{P}$ defines a perfect time τ.

Proof First observe that whenever $g = \widehat{g}_k$ we have $\mathfrak{P} = 1$. Therefore by the above theorem all we need to prove is that τ_0 is strong uniform. This is proved by induction on k. Suppose for each $g \in G$, $l < k$ we have

$$P(\varrho = g | \tau = l) = \frac{1}{|G|}$$

$$P(\tau < k) = |G| \cdot Q^{k-1}(\widehat{g}_{k-1})$$

Then from Proposition 3.3.1 (see also Remark 3.3.2) for any $g \in G$ we get

$$P(X_k = g | \tau \geq k) = Q^k(g) - Q^k(\widehat{g}_{k-1})$$

Therefore

$$P(\varrho = g, \tau = k) = P(X_k = g | \tau \geq k) \cdot \mathfrak{P}(g) = Q^k(\widehat{g}_k) - Q^{k-1}(\widehat{g}_{k-1})$$

From here we immediately get

$$P(\varrho = g | \tau = k) = \frac{1}{|G|}$$

$$P(\tau \leq k) = |G| \cdot Q^k(\widehat{g}_k)$$

and we proved the step of induction. The base of induction is trivial, finishing the proof. $\square$

From Theorem 3.4.2 we immediately get the second part of the Theorem 3.2.7. Note that τ is just one of many possible perfect times. They all however give the same bound on separation distance.

Here is how the construction of τ can be generalized. Suppose we have *any* randomized stopping time τ_0. Denote ϱ_0 its stopping state, $\mathbf{A}_0$ its stopping call, and A_0 the associated algorithm. Consider the following algorithm:

Algorithm G′ Suppose the walk X_t walked along the word $\omega_n \in S_n^{st}$ after i uses of the Algorithm A_0. Flip an unfair coin with probability of heads $\mathfrak{P} = \mathfrak{P}(\omega, i)$.

- if heads, stop.
- if tails, walk until $\mathbf{A} = 1$ (use Algorithm A_0);
- Return to the beginning of the Algorithm.

If τ_0 is simply "walk one step", we get the usual notion of the stopping probability function and Algorithm $\mathbf{G}$. We show that it is possible to define probabilities $\mathfrak{P}(\omega)$ such that Algorithm $\mathbf{G}'$ defines a strong uniform time.

Denote by $Q_i^k(g)$ the probability of a walk being at state $g \in G$ after use of Algorithm A_0 exactly i times which took overall k steps. Since k may vary $Q_i^k(g)$ is no longer a probability distribution on G. Denote by $p_k = p_k(\tau_0)$ the probability of the Algorithm A_0 stopping after k steps. Let $E(\tau_0)$ be the expected number of step of the topping time τ_0. In other words, $E(\tau_0)$ is the mean of the probability distribution p_k.

Call an element $\widehat{g}_{k,i}$ (k,i)-*minimal* if the function $Q_i^k(g)$ minimizes at g:

$$Q_i^k(\widehat{g}_{k,i}) = \min_{g \in G} Q_i^k(g)$$

Definition 3.4.3 For any $g \in G$, $\omega \in S_k^\star(g)$ let

$$\mathfrak{P}(\omega, i) = \mathfrak{P}(g, k, i) = \frac{Q_i^k(\widehat{g}_{k,i}) - \Delta_{k,i}}{Q_i^k(g) - \Delta_{k,i}}$$

where

$$\Delta_{k,i} = \sum_{j=1}^{k-1} Q_{i-1}^j(\widehat{g}_{j,i-1})\, p_{k-j}$$

and $\widehat{g}_{m,l}$ is an (m,l)-minimal element, $m \geq l \geq 1$.

Theorem 3.4.4 The stopping probability function $\mathfrak{P}$ defines a strong uniform time τ.

Example 3.4.5 Consider a "walk one step" stopping time τ_0 (see above). Induction gives

$$\Delta_{k,i} = \delta_{k,i}\, Q^{k-1}(g_{k-1})$$

where $\delta_{k,i}$ is the Kronecker delta. Therefore in this case $\mathfrak{P}$ becomes the the stopping probability function defined in Definition 3.4.1.

Example 3.4.6 Let τ_0 be any strong uniform time for the random walk $\mathcal{W}$. Then for all $g \in G$, $k > 0$

$$Q_1^k(g) = P(\varrho_0 = g | \tau_0 = k) = \frac{1}{|G|}$$

where ϱ_0 is a stopping state of τ_0 (see §3.1). Therefore for all $g \in G$, $k \geq 1$ we get $\mathfrak{P}(g, k, 1) = 1$, i.e. the stopping time τ stops as soon as τ_0 stops. In other words, τ is our old strong uniform time τ_0:

$$\tau = \tau_0$$

Proof of Theorem 3.4.4 The proof is very much similar to the proof of Theorem 3.4.2. We use induction by i and then for each k take the summation over all i, $1 \le i \le k$. $\square$

Sometimes in an attempt to find a strong uniform time one finds a good stopping time τ which is not strong unifrm, but very close in some sense. Here we show that this stoping time can still be used to get bounds on separation.

Theorem 3.4.7 Let $\mathcal{W}$ be a directed random walk, $0 < \epsilon \le 1$. Suppose we are given a randomized stopping time τ_0 such that for every $g \in G$, $k > 0$

$$P(\varrho_0 = g, \tau_0 = k) > \frac{\epsilon}{|G|} P(\tau_0 = k)$$

Then the total separation s for the random walk $\mathcal{W}$ is bounded as

$$s \le \frac{1}{\epsilon} E(\tau_0)$$

As a corollary from the Theorem 3.4.7 we get Theorem 2.6.1. Indeed, simply take τ_0 to be "Walk k steps. Stop".

Proof For each $g \in G$, $i \le k$ from the definition of the stopping probability function $\mathfrak{P}$ we have:

$$\mathfrak{P}(g, k, i) \ge \frac{1}{\epsilon}$$

Since $\mathfrak{P}$ defines a strong uniform time τ, we have:

$$s \le E(\tau) = \frac{1}{\epsilon} E(\tau_0)$$

which proves the result. $\square$

3.5 Time-invariant stopping times.

Definition 3.5.1 Stopping time τ is called *time-invariant* if for all $g \in G$, $k, l > 0$

$$P(\varrho = g | \tau = k) = P(\varrho = g | \tau = l)$$

where ϱ is the stopping state of τ and both conditional distributions are defined. The probability distribution $\mathbf{P}_\tau$ on G defined as

$$\mathbf{p}_\tau(g) = P(\varrho = g | \tau = k)$$

is called *τ-conditional probability distribution*

For example, strong uniform time is a time-invariant stopping time with uniform probability distribution. "Walk k steps" stopping time and hitting times (see §3.6 below) are another extereme examples. Note also that every uniform time-invariant stopping time is strong uniform.

Theorem 3.5.2 Let $\mathcal{W}$ be a directed random walk, $0 < \epsilon \leq 1$. Suppose we are given a time-invariant stopping time τ_0 such that for all $g \in G$

$$\mathbf{p}_\tau(g) \geq \frac{\epsilon}{|G|}$$

Then the total separation s for the random walk $\mathcal{W}$ is bounded as

$$s \leq \frac{1}{\epsilon} E(\tau_0)$$

Proof This is a trivial corollary from Theorem 3.4.7. □

Theorem 3.5.3 Let $\mathcal{W}$ be a directed random walk, τ_0 be a time-invariant stopping time with τ-conditional probability distribution $\mathbf{P}_{\tau_0}$. Suppose random walk $\mathcal{W}_0 = (G, G, \mathbf{P}_{\tau_0})$ is also directed. Then the total separation s of the walk $\mathcal{W}$ is bounded by

$$s \leq E(\tau_0) \cdot s'$$

where s' is the total separation of the random walk $\mathcal{W}_0$.

Note that a priori $\mathcal{W}_0$ does not have to be directed. For example, if τ is defined as follows:

- Walk untill return to e. Stop.

then $\mathbf{P}_{\tau_0}$ is supported on $\{e\}$ and $\mathcal{W}_0$ simply stays at e forever.

Proof of the Theorem Consider a strong uniform time τ defined by its stopping probability function $\mathfrak{P} = \mathfrak{P}(g, k, i)$ in the Definition 3.4.3. Let τ' be a perfect time for the walk $\mathcal{W}_0$ defined by its stopping probability function $\mathfrak{P}_0 = \mathfrak{P}_0(g, i)$ (see Definition 3.4.1).

Since τ_0 is time-invariant, we have

$$Q_i^k(g) = R^i(g) \cdot P(\tau = k)$$

where $R^i(g)$ is the probability of the walk $\mathcal{W}_0$ being at g after i steps. Also one can take $\widehat{g}_{k,i} = \widehat{h}_i$, where $\widehat{h}_i$ is an i-minimal element of the walk $\mathcal{W}_0$. By definition of $\Delta_{k,i}$ for all $k \geq i$ we get

$$\Delta_{k,i} = R^{i-1}(g_{i-1}) \cdot P(\tau = k)$$

and we finally have

$$\mathfrak{P}(g, k, i) = \mathfrak{P}_0(g, i)$$

By Wald's identity (see e.g. [Feller], §18.2, vol 2; [Shiryaev], §7.2) we have

$$s \leq E(\tau) = E(\tau_0) \cdot E(\tau') = E(\tau_0) \cdot s'$$

which proves the result □.

Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk, τ_1, τ_2 - two time-invariant stopping times defined by their stopping probability functions $\mathfrak{P}_1$, $\mathfrak{P}_2$.

Theorem 3.5.4 The stopping time $\tau = \tau_1 + \tau_2$ is time-invariant with τ-conditional probability distribution

$$\mathbf{P}_\tau = \mathbf{P}_{\tau_1} * \mathbf{P}_{\tau_2}$$

Proof Clear. $\square$

Example 3.5.5 Let $G = \mathbb{Z}_4$, $S = \{0, \pm 1\}$, $\mathbf{p}(0) = \frac{1}{2}$, $\mathbf{p}(\pm 1) = \frac{1}{4}$. Define τ_1, τ_2 by the following rules:

- walk till you hit either 1 or -1. Stop.
- walk one step. Stop.

Both τ_1, τ_2 are obviously time-invariant stopping times. The stopping time $\tau = \tau_1 + \tau_2$ was introduced in Example 3.2.4. In Example 3.3.4 we showed that τ is strong uniform. We can use Theorem 3.5.4 to give an independent proof of this result.

We have $\mathbf{p}_{\tau_1}(\pm 1) = \frac{1}{2}$, $\mathbf{p}_{\tau_1}(0) = \mathbf{p}_{\tau_1}(2) = 0$. Also $\mathbf{P}_{\tau_2} = \mathbf{P}$. Therefore $\mathbf{P}_\tau = \mathbf{P}_{\tau_1} * \mathbf{P}_{\tau_2} = U_{\mathbb{Z}_4}$ and τ is indeed strong uniform.

Corollary 3.5.6 Let $G = \mathbb{Z}_p$, p-prime. Suppose a strong uniform time τ can be written as a sum of two time-invariant times $\tau = \tau_1 + \tau_2$. Then one of them is also strong uniform.

Proof By Theorem 3.5.4 we have $U_G = \mathbf{P}_{\tau_1} * \mathbf{P}_{\tau_2}$. The polynomial $1 + z + \cdots + z^{p-1}$ is irreducible in $\mathbb{C}\langle z \rangle / (z^p)$. Therefore either $\mathbf{P}_{\tau_1} = U_G$ or $\mathbf{P}_{\tau_2} = U_G$. Which implies that one of the time-invariant stopping times is strong uniform. $\square$

We finish the section by introducing the notion of the strongly independent stopping times. Although their definition is somewhat hard to understand, they are used crucially in the following sections.

Definition 3.5.7 Stopping times τ_1, τ_2 are called *strongly independent* if they are time-invariant and there exist two generating subsets $S_1, S_2 \subset S$ such that

1) $S = S_1 \cup S_2$, $S_1 \cap S_2 = \{\emptyset\}$, $(\langle S_1 \rangle, \langle S_2 \rangle)$ is an exact decomposition (see §1.1).

2) $s_1 \, s_2 = s_2 \, s_1$ for each $s_1 \in S_1$, $s_2 \in S_2$.

3) $\mathfrak{P}_1(\omega) = 0$ ($\mathfrak{P}_2(\omega) = 0$) for every for every $\omega = w_1 w_2 \dots w_l \in S^\star$ and $w_l \in S_2$ ($w_l \in S_1$).

4) $\mathfrak{P}_1(\omega) = \mathfrak{P}_1(\omega|_{S_1})$ $\big(\mathfrak{P}_2(\omega) = \mathfrak{P}_2(\omega|_{S_2})\big)$ for every $\omega = w_1 w_2 \dots w_l \in S^\star$ and $w_l \in S_1$ ($w_l \in S_2$).

5) $P_{\tau_1} = U_{\langle S_1 \rangle}$, $P_{\tau_2} = U_{\langle S_2 \rangle}$.

Observe that condition 5) means that the stopping times τ_1, τ_2 are time-invariant with τ-conditional probability distribution uniform on a subgroup generated by S_1, S_2.

The canonical example for the strongly independent stopping times is the case when we have a direct product of of two groups $G = G_1 \times G_2$ and strong uniform

times for random walks on each of the factors. Namely, suppose we are given random walks $\mathcal{W}_1 = (G_1, S_1, \mathbf{P}_1)$ and $\mathcal{W}_2 = (G_2, S_2, \mathbf{P}_2)$. Let τ_1', τ_2' be the strong uniform times for each of the walks. Define a random walk $\mathcal{W}$ by choosing in random a factor G_1 or G_2 and making a step of the walk $\mathcal{W}_1$ or $\mathcal{W}_2$ on that subgroup. Note that the first condition then simply indicates that this random walk is generated by a union of two generating sets.

Consider the following stopping times τ_1, τ_2. Let τ_1 (τ_2) observe movements of the walk only on G_1 (G_2) and ignore any steps made on the other factor G_2 (G_1). We want τ_1 (τ_2) to stop only when they would have stopped if it was just a walk on G_1 (G_2). This is basically a restatement of the condition 4) in this case. The condition 3) basically means that stopping times are allowed to stop only in their turn, i.e. when the step is made on a corresponding factor. The condition 5) together with time-inariance means that τ_1, τ_2 are strong uniform.

The main idea behind stating this general definition rather than stating the example above is to take only those properties of the stopping times τ_1, tau_2 that are necessary for proving the Theorem below. In §4.2 we consider other examples of the strongly independent stopping times.

Example 3.5.8 Let $G = \mathbb{Z}_4$, $S = \{e_1, e_2, \pm 1\}$, $\mathbf{p}(e_{1,2}) = \mathbf{p}(\pm 1) = \frac{1}{4}$, where $e_1 = e_2 = 0$ are formally different identity elements. This means that when applying Procedure **W** (see §2.1) we get different words $\omega_{t+1} = \omega_t \cdot e_1$ or $\omega_{t+1} = \omega_t \cdot e_2$, but the walk actually stays in both cases. Observe that the random walk $\mathcal{W} = (G, S, \mathbf{P})$ is equivalent to the random walk introduced in Example 2.1.1.

Define τ_1, τ_2 by the following rules:

- walk till either e_1 or 1 is used. Stop.
- walk till either e_2 or -1 is used. Stop.

We claim the τ_1 and τ_2 are strongly independent time-invariant stopping times. Indeed, take $S_1 = \{e_1, 1\}$, $S_1 = \{e_2, -1\}$. The only property to check is 5) which follows immediately from the definition of **P**.

Theorem 3.5.9 Suppose τ_1, τ_2 are strongly independent time-invariant stopping times. Then $\tau = \max(\tau_1, \tau_2)$ is a strong uniform time.

Proof Compute the conditional probability of stopping at a given state. Let ϱ, ϱ_1 and ϱ_2 be the stopping states of the stopping times τ, τ_1 and τ_2 respectively.

First recall that by 1) we have an exact decomposition. Denote $p = \sum_{s \in S_1} \mathbf{p}(s)$. Let $g \in G$, $g = g_1\, g_2$ be the decomposition of g, $g_1 \in \langle S_1 \rangle$, $g_2 \in \langle S_2 \rangle$. By $\mathfrak{P}(\omega)$, $\omega \in S^\star$ we denote the stopping probability function for the associated with the stopping time $\tau = max(\tau_1, \tau_2)$ (see §3.1). By $\mathbf{A}_1$, $\mathbf{A}_2$ we denote stopping calls of τ_1, τ_2.

Recall the interpretation of the random walk in terms of words in the language $S^\star$ (see §2.1). We have:

$$P(\varrho = g,\, \tau = k) \;=\; \sum_{\omega = s_1 \dots s_k \in S_k^\star(g)} \mathbf{p}(s_1) \cdot \ldots \cdot \mathbf{p}(s_k) \cdot \mathfrak{P}(\omega)$$

Since $\tau = \max(\tau_1, \tau_2)$ we can use properties 3), 4) to break the summation into two depending which stopping time finishes first. Also break the sum on the right

hand side into summation over pairs of subwords in the language $(S_1)^\star$ and $(S_2)^\star$. We separate the last s_k from the rest and take a summation over all l which denotes the length of the first subword (cf. Theorem 2.1.5).

$$= \sum_{l=0}^{k-1} \binom{k-1}{l} p^l (1-p)^{k-l-1} \left(p \cdot \sum_{\substack{\omega_1 = s_{i_1} \dots s_{i_l} \\ s_{i_1}, \dots, s_{i_l} \in S_1 \\ \gamma(\omega_1) = g_1,\ \mathbf{A}_1(\omega_1) = 1}} \mathbf{p}(s_{i_1}) \cdot \ldots \cdot \mathbf{p}(s_{i_l}) \times \right.$$

$$\times \sum_{\substack{\omega_2 = s_{j_1} \dots s_{j_{k-l-1}},\ \ s_k \in S_2 \\ s_{j_1}, \dots, s_{j_{k-l-1}} \in S_2 \\ \gamma(\omega_2) \cdot s_k = g_2,\ \mathbf{A}_2(\omega_2) = 0}} \mathbf{p}(s_{j_1}) \cdot \ldots \cdot \mathbf{p}(s_{j_{k-l-1}}) \cdot \mathbf{p}(s_k) \cdot \mathfrak{P}_2(\omega_2 s_k) +$$

$$+ (1-p) \sum_{\substack{\omega_1 = s_{i_1} \dots s_{i_l},\ s_k \in S_1 \\ s_{i_1}, \dots, s_{i_l} \in S_1 \\ \gamma(\omega_1) \cdot s_k = g_1,\ \mathbf{A}_1(\omega_1) = 0}} \mathbf{p}(s_{i_1}) \cdot \ldots \cdot \mathbf{p}(s_{i_l}) \cdot \mathbf{p}(s_k) \cdot \mathfrak{P}_1(\omega_1 s_k) \times$$

$$\left. \times \sum_{\substack{\omega_2 = s_{j_1} \dots s_{j_{k-l-1}} \\ s_{j_1}, \dots, s_{j_{k-l-1}} \in S_2 \\ \gamma(\omega_2) = g_2,\ \mathbf{A}_2(\omega_2) = 1}} \mathbf{p}(s_{j_1}) \cdot \ldots \cdot \mathbf{p}(s_{j_{k-l-1}}) \right)$$

Now use the property 5) of τ_1, τ_2 being uniform on $\langle S_1 \rangle$, $\langle S_2 \rangle$ and time-invariant:

$$= \sum_{l=0}^{k-1} \binom{k-1}{l} p^l (1-p)^{k-l-1} \left(p \cdot \frac{P(\tau_1 \le l)}{|\langle S_1 \rangle|} \cdot \frac{P(\tau_2 = k-l)}{|\langle S_2 \rangle|} + \right.$$

$$\left. + (1-p) \cdot \frac{P(\tau_1 = l+1)}{|\langle S_1 \rangle|} \cdot \frac{P(\tau_2 \le k-l-1)}{|\langle S_2 \rangle|} \right)$$

Now observe that the last equation does not depend on $g = g_1 \cdot g_2$. Therefore for all $g, g' \in G$ we have

$$P(\varrho = g,\, \tau = k) = P(\varrho = g',\, \tau = k)$$

Thus τ is strong uniform. It finishes the proof of the theorem. $\square$

We use the notion of strongly independent stopping times and Theorem 3.5.9 in §4.3 when discussing random walks on a direct products of groups.

Example 3.5.10 Let $\mathcal{W} = (G, S, \mathbf{P})$, τ_1, τ_2 be as in previous example. The stopping time τ can be defined by the following rule:

- walk till either e_1 or 1 and either e_2 or -1 is used. Stop.

By Theorem 3.5.9 τ is a strong uniform time. It is also a perfect time since once we get to 2 we stop (cf. Examples 3.3.4, 3.5.5).

3.6 Hitting times.

Recall the definitions in 2.2*b*. By a *g-hitting time* ht_g we mean the expected time to get to $g \in G$ for the first time. Let $ht_e = 0$. The *hitting time* ht and the *average hitting time* aht are defined as follows:

$$ht = \max_{g \in G} ht_g$$

$$aht = \frac{1}{|G|} \sum_{g \in G} ht_g$$

The cover time ct is define as an expected time to hit all the elements.

Theorem 3.6.1 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk, s the total separation of $\mathcal{W}$. Then

1) $s + aht \geq ht$

2) $s + aht = ht$ if $\mathcal{W}$ has an extremal element $\widehat{g} \in G$ and $ht = ht_{\widehat{g}}$

Proof Let $\widetilde{\tau}$ be a perfect time for the random walk $\mathcal{W}$. Suppose $ht = ht_g$. Let τ_g be the stopping time defined as follows:

• walk till you hit g. Stop.

Observe that τ_g is time-invariant. Define $\tau = \widetilde{\tau} + \tau_g$. Since this stops only at g it is also time-invariant and

$$E(\tau) \geq E(\tau_g)$$

Since $E(\tau) = E(\widetilde{\tau}) + aht = s + aht$ we have

$$s + aht = E(\tau) \geq E(\tau_g) = ht$$

which proves the first part of the Theorem.

To prove the second part, simply observe that τ always stops at $\widehat{g}$ whenever it gets there. Indeed, the perfect time $\widetilde{\tau}$ respects $\widehat{g}$ by Theorem 3.3.3 and since $\widehat{g}$ is an extremal element for $\mathcal{W}$. Also τ_g always stops at $\widehat{g}$ by definition. Therefore the stopping times τ and τ_g are identical:

$$\tau = \tau_g$$

and $s + aht = E(\tau) = E(\tau_g) = ht$. □

Example 3.6.2 For the random walk $\mathcal{W}$ on $\mathbb{Z}_4$ (see Examples 2.2.13, 2.4.6) we have computed

$$ht = 8\,, \quad aht = 5$$

Observe that element 2 is an extremal element. Also $ht = ht_2$. Therefore by Theorem 3.6.1 we have

$$s = ht - aht = 3$$

which agrees with our previous computations.

Theorem 3.6.3 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk, s the total separation of $\mathcal{W}$. Then

$$s \cdot |G| \geq ht$$

$$s \cdot |G| \cdot \mathfrak{h}_{|G|} \geq ct$$

Proof Let $\widetilde{\tau}$ be a perfect time for the random walk $\mathcal{W}$. Let $\mathbf{A}$ be its stopping call. Suppose $ht = ht_g$. Consider the following stopping time τ_1:

• walk till $\mathbf{A} = 1$. If the walk is not at g, start all over. Otherwise stop.

Clearly the expected time $E(\tau_1) \geq ht$. On the other hand, by Wald's identity (see e.g. [Feller], §18.2, vol 2; [Shiryaev], §7.2) we have:

$$E(\tau_1) = E(\widetilde{\tau}) \cdot |G| = s \cdot |G|$$

This proves the first part.

The proof of the second part is analogous. Consider a stopping time τ_2:

• walk till $\mathbf{A} = 1$. If the walk hit all the elements, stop. Otherwise start all over.

Obviously $E(\tau_2) \geq ct$. On the other hand, the expected number of steps to hit all the elements if we were picking elements of G in random is equal to $|G| \cdot \mathfrak{h}_{|G|}$. Thus by the Wald identity we have:

$$ct \leq E(\tau_2) \leq E(\widetilde{\tau}) \cdot |G| \cdot \mathfrak{h}_{|G|} = s \cdot |G| \cdot \mathfrak{h}_{|G|}$$

which proves the second part of the Theorem. $\square$.

Example 3.6.4 Let $\mathcal{W}$ be as in the previous example. We have $|G| = 4$, $\mathfrak{h}_{|G|} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = 2\frac{1}{12}$. Therefore Theorem 3.6.3 gives

$$s \cdot 4 \geq ht \ , \quad s \cdot 8\frac{1}{3} \geq ct$$

which agrees with our previously computed values $s = 3$, $ht = 8$ and $ct = 12$.

We can think of theorems 3.6.1, 3.6.3 as lower bounds on the total separation in terms of hitting and cover times. Let us generalize the first part of the Theorem 3.6.3 in the following way.

Theorem 3.6.5 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk, s the total separation of $\mathcal{W}$. Suppose $B \subset G$ is a subset of G, ht_B is an expected time for a random walk $\mathcal{W}$ to hit any element in B. Then

$$s \geq \frac{|B|}{|G|}\, ht_B$$

Proof Let $\widetilde{\tau}$ be a perfect time for the random walk $\mathcal{W}$, $\mathbf{A}$ is its stopping call. Consider the following stopping time τ_1:

• walk till $\mathbf{A} = 1$. If the walk is not at B, start all over. Otherwise stop.

By analogy with the proof of the Theorem 3.6.3 we have

$$E(\tau_1) = \frac{|G|}{|B|} E(\tau) = \frac{|G|}{|B|}\, s$$

On the other hand, $E(\tau_1) \geq ht_B$. This proves the result. $\square$

Example 3.6.6 Let $G = U(3; \mathbb{F}_2)$ (see Example 2.6.5), $S = \{e, R_1, R_2\}$, $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(R_1) = \mathbf{p}(R_2) = \frac{1}{4}$. Consider a random walk $\mathcal{W}(G, S, \mathbf{P})$.

Denote by B the set of upper triangular matrices $M \in G$ with 1 in the upper right corner. The expected time to hit B is the expected time before we apply R_1 and then the expected time before we apply R_2. Therefore

$$s \geq \frac{|B|}{|G|}(4+4) = \frac{4}{8} \cdot 8 = 4$$

which we already know since $d_S = 4$.

Example 3.6.7 Let $G = \mathbb{Z}_m$, $\mathcal{W}$ be a standard random walk on $\mathbb{Z}_m$. Suppose $m = 4\,k + 2$. Define $B = \{k+1, \ldots, 3\,k+1\}$. It is known that

$$ht_B = ht_{\{1 \pm k\}} = 2\,(k+1)^2$$

(see e.g. [Feller], vol 1, §14.3). From here we have

$$s \geq \frac{2\,k+1}{4\,k+2}\, 2\,(k+1)^2 > \frac{1}{16}\, m^2$$

Example 3.6.7 Let $G = \mathbb{Z}_2^n$, n even. Let $\mathcal{W}$ be a standard random walk on $\mathbb{Z}_2^n$ (see Example 2.5.2). We can think about this walk as the nearest neighbor walk on a cube. Denote $d(x, y)$ be a Hamming distance between two points (the number of different coordinates). Let B be a Hamming ball around $(1, \ldots, 1)$ of radius $n/2$. The hitting time ht_B is known :

$$ht_B = \frac{1}{2}\,(n \ln(n) + O(1))$$

(see e.g. [Matt2]). From Theorem 6.3.5 we immediately get

$$s \geq \frac{|B|}{|G|} \cdot ht_B \sim \frac{1}{4}\, n \ln(n)$$

In §4.3 we show that this lower bound is of the right order.

Remark 3.6.8 In [A2] Aldous proved the lower bound for the cover time:

$$ct \geq C \cdot |G| \cdot \log(|G|)$$

for the cover time of random walks. When the walk is rapidly mixing, i.e. the total separation s is relatively small compared to $|G|$ Theorem 3.6.3 gives good bounds on the cover time. For example, if $G = \mathbb{Z}_2^n$ and $\mathcal{W}$ is the standard random walk on $\mathbb{Z}_2^n$ we have $s \sim n \ln(n)$ (see §4.2) which gives us

$$C_1\, n\, 2^n \leq ct \leq C_2 n^2\, \log(n)\, 2^n$$

The right bound in this case is due to Matthews: $ct \sim n\, 2^n$ (see [Matt2]).

4. Basic constructions

4.1 Cyclic groups.

Here we are present a construction of a strong uniform time for a random walk on a $\mathbb{Z}_n$. As in §2, we take a set of generators $S = \{-1, 0, +1\}$ and a probability distribution $\mathbf{P}(0) = \frac{1}{2}$ and $\mathbf{P}(-1) = \mathbf{P}(+1) = \frac{1}{4}$.

First, generalize the construction given in Example 3.2.5 for the case $n = 4$. Let $n = 2^m$.

Algorithm 4.1.1

- Walk till you hit either 2^{m-2} or $3 \cdot 2^{m-2}$.
- Walk till you hit either 2^{m-3}, $3 \cdot 2^{m-3}$, $5 \cdot 2^{m-3}$ or $7 \cdot 2^{m-3}$.
- ...
- Walk till you hit either 1, 3, 5, ..., $2^m - 1$.
- Do one more step. Stop.

Theorem 4.1.2 The Algorithm 4.1.1 defines a perfect time τ.

First Proof First, observe that τ is a usual (not a randomized) stopping time. Denote $\mathcal{X}_n$ the set of walk paths on $\mathbb{Z}_n$. Let $I(l, m, n)$, $1 \le l < m$ be the set of sequences $(i) = (i_1, i_2, \ldots, i_l)$, $i_1, \ldots, i_{l-1} \in \mathbb{Z}_n$ such that

$$\begin{aligned} & i_1 \in \{\pm 2^{m-2}\}, \\ & i_2 \in \{i_1 \pm 2^{m-3}\}, \\ & \ldots \\ & i_l \in \{i_{l-1} \pm 2^{m-l-1}\} \end{aligned}$$

Define $\tau_{l,m}(X)$, $X \in \mathcal{X}_n$ to be the minimal k such that $X \supset (i)$, i.e. a sequence $(X_1, X_2, \ldots, X_k)$ contains a subsequence $(i_1, i_2, \ldots, i_l) \in I(l, m, n)$. By symmetry $(i_v - i_{v+1})$, $v = 1, \ldots, l-1$ can occur in $(i_1, i_2, \ldots, i_l)$ with equal probability and these are independent events. Therefore for any $(i) \in I(l, m, n)$

$$P(X \subset (i) | \tau_{l,m}(X) = k) = \frac{1}{|I(l, m, n)|} = \frac{1}{2^l}$$

When $n = 2^m$ the set of sequences $I(m-1, m, n)$ is in a one-to-one correspondence with its last element $i_{m-1} = 1 \pmod 2$. From here we get

$$P(X_k = 2\,j - 1 | \tau_{m-1,m} = k) = \frac{1}{2^{m-1}}$$

for all $1 \le j \le 2^{m-1}$, $k > 0$.

Observe that if $n = 2^m$, $\tau = \tau_{m,m-1} + 1$. Therefore for all $1 \le j \le 2^{m-1}$, $k > 0$ we have

$$P(\varrho = 2\,j - 1 \mid \tau = k) = \frac{\mathbf{p}(0)}{2^{m-1}} = \frac{1}{2^m}$$
$$P(\varrho = 2\,j \mid \tau = k) = \frac{\mathbf{p}(1) + \mathbf{p}(-1)}{2^{m-1}} = \frac{1}{2^m}$$

and we have proved that τ is strong uniform.

We claim that $n/2 = 2^{m-1}$ is an extremal element (see §3.3). Simply observe that in order for the walk to reach 2^{m-1} it must first pass $i_1 = \pm 2^{m-2}$, then $i_2 = \pm 3 \cdot 2^{m-3}$, $\dots$, $i_{m-1} = 2^{m-1} \pm 1$. This means that according to the algorithm whenever the walk reaches $n/2$ it stops there. Since τ is strong uniform, we have proved both that $n/2$ is an extremal element and that τ is perfect. $\square$

Second Proof

Define stopping times τ_i, $1 \le i \le m-1$ by the following rule:

• walk till you hit $\pm 2^{m-1-i}$. Stop.

For the stopping time τ_0:

• walk one step. Stop.

Obviously, these stopping times are time-invariant and thus so is $\tau = \tau_1 + \dots + \tau_{m-1} + \tau_0$. Compute $\mathbf{P}_\tau$. Fix a, $a^n = 1$. We have

$$(a^{2^{m-2}} + a^{-2^{m-2}}) \cdot \ldots \cdot (a + a^{-1}) \cdot (2 + a + a^{-1})$$
$$= a^{-n/2-1}\,(a-1)\,(1 + a + \dots + a^{n-1}) = (1 + a + \dots + a^{n-1})$$

Therefore by Theorem 3.5.4 we get

$$\mathbf{P}_\tau = \mathbf{P}_{\tau_1} * \dots * \mathbf{P}_{\tau_{m-1}} * \mathbf{P}_{\tau_0} = U_G$$

and we proved that τ is strong uniform. $\square$

Corollary 4.1.3 Let s be the total separation for the standard random walk $\mathcal{W}$ on $\mathbb{Z}_n$, $\mathcal{W} = (\mathbb{Z}_n, \{0, \pm 1\}, \mathbf{P})$, $n = 2^m$. Then

$$s = \frac{1}{6}\,n^2 + \frac{1}{3}$$

Example 4.1.4 When $n = 2, 4, 8$ we get $s = 1, 3, 11$ (cf. Examples 2.2.12, 2.2.13, 2.6.5).

Proof Denote E_z the expected time for the walk $\mathcal{W}$ to hit $\pm z$, $z < \frac{n}{2}$. Finding E_z is equivalent to solving the famous bankruptcy problem (see e.g. [Feller], vol 1, §14.3) which gives us

$$E_z = 2\,z^2$$

In the notation of the second proof, $E(\tau_i) = E_{2^{m-i-1}}$, $1 \le i \le m-1$. Since τ is a perfect time, we have

$$s = E(\tau) = E_{2^{m-2}} + E_{2^{m-3}} + \dots + E_1 + 1 = 2\,\left(4^{m-2} + 4^{m-3} + \dots + 1\right) + 1$$
$$= 2\,\frac{4^{m-1} - 1}{4 - 1} + 1 = \frac{1}{6}\,n^2 + \frac{1}{3}$$

This proves the Corollary. $\square$

Now suppose n is not a power of 2. It is not hard to see that the element $a = \lfloor \frac{n}{2} \rfloor$ is an extremal element, so theoretically we should be able to compute the total separation by use of probabilities $Q^k(a)$:

$$s = 1 + \sum_{k=1}^{\infty} (1 - Q^k(a))$$

Unfortunately computing $Q^k(a)$ is not an easy task so we use another approach.

Theorem 4.1.5 Let s be the total separation for the standard random walk $\mathcal{W}$ on $\mathbb{Z}_n$, $\mathcal{W} = (\mathbb{Z}_n, \{0, \pm 1\}, \mathbf{P})$. Then

$$s \leq \frac{4}{3}\, n^2 + \frac{2}{3}$$

Proof Observe that the stopping times $\tau_{l,m}$ in the proof of the Theorem 4.1.2 are defined for any n. Let m be an integer such that $2^{m-1} < n \leq 2^m$. Denote $\tau_0 = \tau_{m-1,m}$, ϱ_0 - the stopping state of τ_0. For any $g \in G$, $k > 0$ we have

$$P(\varrho_0 = g | \tau_0 = k) \geq \frac{1}{2^m}$$

Thus τ_0 is time-invariant and

$$P_{\tau_0} \geq \frac{n}{2^m} \cdot \frac{1}{|\mathbb{Z}_n|}$$

From Theorem 3.4.7 we get

$$s \leq \frac{2^m}{n}\, E(\tau_0) = \frac{2^m}{n} \cdot \frac{2^{2\,m} + 2}{6} < \frac{4\, n^2 + 2}{3}$$

which finishes the proof. $\square$

Remark 4.1.6 The Theorem 4.1.2, Corollary 4.1.3 and the Theorem 4.1.5 with a slightly different proofs were obtained by Diaconis and Fill in [DF].

Theorem 4.1.7 Let s be the total separation for the standard random walk $\mathcal{W}$ on $\mathbb{Z}_n$. Then

$$s = \frac{1}{6}\, n^2 + \frac{1}{3}, \quad n - \text{even}$$
$$s = \frac{1}{6}\, n^2 + \frac{1}{6}, \quad n - \text{odd}$$

Proof Let E_i be the expected time to hit i, $0 \leq i \leq n-1$. It is known that $E_i = 2\, i\, (n - i)$ (see e.g. [Feller], vol 1, §14.3). Therefore

$$aht = \frac{1}{n} \sum_{i=0}^{n-1} 2\, i\, (n - i) = \frac{1}{3}\, n^2 - \frac{1}{3}$$

Let $m = \lfloor \frac{n}{2} \rfloor$. Then

$$ht = ht_m = 2\,m\,(n-m)$$

Recall that m is also an extremal element of $\mathcal{W}$ (see e.g. [D], §3D, Exercise 10). By part 2) of Theorem 3.6.1 we get

$$s = ht - aht = \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil - \frac{1}{3}\,n^2 + \frac{1}{3}$$

Considering separately two cases when n is even and odd, we get the result. □

4.2 Change of $\mathbf{p}(e)$.

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$ with $\mathbf{p}(e) = p$, $0 < p < 1$. There is a natural way to rescale the probabilities of generators in order to change the probability of identity $\mathbf{p}(e)$. Indeed, for any p', $0 \le p' < 1$ define a random walk $\mathcal{W}' = (G, S, \mathbf{P}')$ as follows:

$$\mathbf{p}'(e) = p'$$
$$\mathbf{p}'(s) = \mathbf{p}(s)\frac{1-p'}{1-p}\ ,\ s \in \check{S}$$

We say that the random walk $\mathcal{W}'$ is the *rescaled* random walk $\mathcal{W}$.

We claim that when $p < p'$ there is an easy bound on the separation distance s'_k for the walk $\mathcal{W}'$ in terms of the separation distance s_k for $\mathcal{W}$.

Theorem 4.2.1 Let s, s' be the total separations for the random walks $\mathcal{W} = (G, S, \mathbf{P})$, $\mathcal{W}' = (G, S, \mathbf{P}')$ defined as above. If $p \le p'$, then

$$s' \le \frac{1-p}{1-p'} \cdot s$$

Moreover, if $\mathcal{W}$ has an extremal element $\widehat{g}$, then $\mathcal{W}'$ has an extremal element $\widehat{g}$ and

$$s' = \frac{1-p}{1-p'} \cdot s$$

Corollary 4.2.2 Under the conditions of Theorem 4.2.1, if $p' = \frac{1}{2}$, then

$$s' \le 2\,s$$

Proof Clear.□

Corollary 4.2.3 Let $\mathcal{W} = (G, G, \mathbf{P})$ be a uniform random walk with a generating set G. If $\mathbf{p}(e) \ge \frac{1}{|G|}$, then

$$s = \frac{1 - \frac{1}{|G|}}{1 - \mathbf{p}(e)}$$

Proof Clear. □

Proof of the Theorem We are going to prove Theorem 4.2.1 by constructing a strong uniform time for the random walk $\mathcal{W}'$.

Let τ be a perfect time for the random walk $\mathcal{W}$. This exists by Theorem 3.2.7. Suppose our stopping time τ is defined in terms of Algorithm **G** (see §3.1) with stopping probability function $\mathfrak{P}$. Define a new stopping time by the following algorithm.

Let $S_1 = S \cup e_1$ be a new set of generators with two formally different identity generators e, e_1. This means that no matter which of these two generators is chosen in Procedure **W** (see §2.1), the walks stays although we get two different words in $(S_1)^\star$. Since $p' \geq p$ we can consider the following probability distribution $\mathbf{P}_1$:

$$\begin{aligned} \mathbf{p}_1(e_1) &= p' - p \cdot \frac{1-p'}{1-p} \\ \mathbf{p}_1(e) &= p \cdot \frac{1-p'}{1-p} \\ \mathbf{p}_1(s) &= \mathbf{p}(s) \cdot \frac{1-p'}{1-p} \ , \ s \in \check{S} \end{aligned}$$

Observe that the random walk $\mathcal{W}_1 = (G, S_1, \mathbf{P}_1)$ is equivalent to $\mathcal{W}'$. Therefore rather than looking at $\mathcal{W}'$ we are going to study $\mathcal{W}_1$.

For a word $\omega \in S_1^\star$ define $\omega|_S \in S^\star$ to be its S-subword, i.e a subword obtained from ω after erasing all the letters e_1.

Define a new stopping probability function $\mathfrak{P}_1$ as follows

$$\mathfrak{P}_1(\omega) = \begin{cases} 0, \, w_n = e_1 \\ \mathfrak{P}(\omega|_S), \, w_n \in S \end{cases}$$

We claim that the Algorithm **G** with stopping probability function $\mathfrak{P}_1$ defines a strong uniform time τ_1. Indeed, by definition of $\mathfrak{P}_1$ and Theorem 2.1.5 applied to the identity element e_1, we have:

$$\begin{aligned} P(\varrho_1 = g | \tau_1 = k) &= \sum_{\omega = s_1 \dots s_k \in S_1^\star} \mathbf{p}_1(s_1) \cdot \ldots \cdot \mathbf{p}_1(s_k) \cdot \mathfrak{P}_1(\omega) \\ &= \sum_{l=1}^{k} (\mathbf{p}_1(e_1))^{k-l} \, (1 - \mathbf{p}_1(e_1))^{l-1} \binom{k-1}{l} \sum_{\omega' = s'_1 \dots s'_l \in S^\star} \mathbf{p}_1(s'_1) \cdot \ldots \cdot \mathbf{p}_1(s'_k) \cdot \mathfrak{P}(\omega') \\ &= \sum_{l=1}^{k} (\mathbf{p}_1(e_1))^{k-l} \, (1 - \mathbf{p}_1(e_1))^{l-1} \binom{k-1}{l} P(\varrho = g | \tau = l) \\ &= \sum_{l=1}^{k} (\mathbf{p}_1(e_1))^{k-l} \, (1 - \mathbf{p}_1(e_1))^{l} \binom{k-1}{l} \frac{1}{|G|} \quad = \quad \frac{1}{|G|} \end{aligned}$$

where $\varrho_1 = X_{\tau_1}$ is a stopping state. We have a binomial coefficient $\binom{k-1}{l}$ in the second equality since whenever $s_k = e_1$, $\mathfrak{P}_1(\omega) = \mathfrak{P}_1(s_1 \dots s_k) = 0$. The fourth equality follows from τ being a strong uniform time.

This proves that τ_1 is a strong uniform time. Now we need to compute the expected stopping time of τ_1.

Denote $p_1 = \mathbf{p}_1(e_1) = \frac{p'-p}{1-p}$. Observe that expected number of times the element e_1 occurs between two elements $s_i, s_{i+1} \in S$ in a word $\omega \in S_1^\star$ is equal to

$$E_0 = p_1\,(1-p_1) + 2\,p_1^2\,(1-p_1) + 3\,p_1^3\,(1-p_1) + \cdots = \frac{p_1}{1-p_1} = \frac{p'-p}{1-p'}$$

Use Wald's identity to get the expected stopping time of τ_1 (see e.g. [Feller], §18.2, vol 2; [Shiryaev], §7.2)

$$E(\tau_1) = (1+E_0)\,E(\tau)$$

Since τ_1 is a strong uniform time and τ is a perfect time, by the Theorem 3.2.7 we have

$$s' \le E(\tau_1) = \frac{1-p}{1-p'} \cdot E(\tau) = \frac{1-p}{1-p'} \cdot s$$

This finishes proof of the first part of Theorem 4.2.1.

Suppose now that the random walk $\mathcal{W}$ has an extremal element $\widehat{g}$. By Theorem 3.3.3 the perfect time τ must respect $\widehat{g}$. We show that τ_1 also respects $\widehat{g}$. Indeed, by definition of the stopping probability function $\mathfrak{P}_1$, the first time we get to the state $\widehat{g}$ we have $\mathfrak{P}_1(\omega) = \mathfrak{P}(\omega') = 1$. This means that τ_1 respects which implies that it is also perfect. Therefore

$$s' = E(\tau_1) = \frac{1-p}{1-p'} \cdot s$$

which finishes the proof of the second part of the Theorem 4.2.1 $\square$

It is possible to define a stopping time τ' for the random walk $\mathcal{W}'$ without consideration of a new letter e_1. This idea, however, is very useful and will be applied later in several other examples.

Example 4.2.4 As in Example 2.2.12, consider the group $G \simeq \mathbb{Z}_2$ with $S = G = \{e, a\}$. Define two probability distributions $\mathbf{P}$ and $\mathbf{P}'$ as follows:

$$\mathbf{p}(a) = \frac{p}{2},\ \mathbf{p}(e) = 1 - \frac{p}{2},\ p < 1$$
$$\mathbf{p}'(a) = \frac{p'}{2},\ \mathbf{p}'(e) = 1 - \frac{p'}{2},\ p' < p$$

For these random walks we get (see Example 2.2.12)

$$s = \frac{1}{p}, \quad s' = \frac{1}{p'}$$

Since both walks have an extremal element at 1, the theorem above gives us

$$s' = \frac{1-\left(1-\frac{p}{2}\right)}{1-\left(1-\frac{p'}{2}\right)} \cdot s = \frac{p}{p'} \cdot \frac{1}{p} = \frac{1}{p'}$$

which agrees with our previous computations.

4.3 Direct product construction.

Suppose we are given two finite groups G_1 and G_2 and two random walks $\mathcal{W}_1 = (G_1, S_1, \mathbf{P}_1)$, and $\mathcal{W}_2 = (G_2, S_2, \mathbf{P}_2)$. There is a natural way to construct a random walk $\mathcal{W} = (G, S, \mathbf{P})$ on a direct product of the two groups:

$$G = G_1 \times G_2$$
$$S = S_1 \times \{e_2\} \cup \{e_1\} \times S_2 \in G$$

$$\mathbf{p}(s', e_2) = \frac{1}{2}\,\mathbf{p}_1(s'),\ \ s' \in \check{S}_1$$
$$\mathbf{p}(e_1, s'') = \frac{1}{2}\,\mathbf{p}_2(s''),\ \ s'' \in \check{S}_2$$
$$\mathbf{p}(e_1, e_2) = \frac{1}{2}\,(\mathbf{p}_1(e_1) + \mathbf{p}_2(e_2))$$

We call this random walk $\mathcal{W}$ a *direct product* of the random walks $\mathcal{W}_1$ and $\mathcal{W}_2$:

$$\mathcal{W} = \mathcal{W}_1 \times \mathcal{W}_2$$

Rather than taking weights $\left(\frac{1}{2}, \frac{1}{2}\right)$ for the two random walks $\mathcal{W}_1$ and $\mathcal{W}_2$, one can consider a $(p, 1-p)$-*weighted direct product* $\mathcal{W}$ defined analogously $\mathcal{W}_1 \times_{(p,1-p)} \mathcal{W}_2 = (G, S, \mathbf{P}_{(p,1-p)})$, where

$$\mathbf{p}_{(p,1-p)}(s', e_2) = p\,\mathbf{p}_1(s'),\ \ s' \in \check{S}_1$$
$$\mathbf{p}_{(p,1-p)}(e_1, s'') = (1-p)\,\mathbf{p}_2(s''),\ \ s'' \in \check{S}_2$$
$$\mathbf{p}_{(p,1-p)}(e_1, e_2) = p\,\mathbf{p}_1(e_1) + (1-p)\,\mathbf{p}_2(e_2))$$

In order to get a uniform random walk out two uniform random walks one should take $p = \frac{|S_1|}{|S_1|+|S_2|}$ (see definition of a uniform random walk in §2.1). It is also convenient to consider a $(p_1, p_2, \ldots, p_n)$-weighted direct product of n random walks $\mathcal{W}_1,\ \mathcal{W}_2,\ \ldots,\ \mathcal{W}_n$, where $p_1 + p_2 + \cdots + p_n = 1$. They can be defined analogously. By a *n-th power* of a walk $\mathcal{W}$ we mean a $\left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$-weighted direct product $\mathcal{W}^n = \mathcal{W} \times \cdots \times \mathcal{W}$ (n times).

For simplicity we consider the usual directed products of random walks as well as their powers.

Example 4.3.1 Consider a standard random walk on a cyclic group $\mathbb{Z}_4 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$ (see Example 2.2.13). It is easy to check that this walk is a product of two random walks on $\mathbb{Z}_2$.

Example 4.3.2 Consider a standard random walk on an n-cube $\mathbb{Z}_2^n$ (see Example 2.5.2) with a set of generators $S = \{e, (0, 0, \ldots, 1, \ldots, 0)\}$, $|S| = n$ and probability distribution $\mathbf{P}$ defined as follows:

$$\mathbf{p}(e) = \frac{1}{2}, \qquad \mathbf{p}(s) = \frac{1}{2\,n}\,, \quad s \in \check{S}$$

One can show that this random walk is a n-th power of a random walk on $\mathbb{Z}_2$.

We can think of $\mathcal{W}$ as two separate random walks on G_1 and G_2 with steps randomly taken either in one group or the other.

Now suppose we are given strong uniform times τ_1, τ_2 with stopping probability functions $\mathfrak{P}_1$, $\mathfrak{P}_2$ defined for the random walks $\mathcal{W}_1$ and $\mathcal{W}_2$ respectively. We will present a strong uniform time τ for the direct product $\mathcal{W} = \mathcal{W}_1 \times_{(p,1-p)} \mathcal{W}_2$. Here is an informal way to do it.

Algorithm 4.3.3

• walk till both rules stop the corresponding random walks on G_1 and G_2. Stop.

Formally, define the time-invariant stopping times τ_1', τ_2' by the stopping probability function $\mathfrak{P}_1'$, $\mathfrak{P}_1'$ as follows:

$$\mathfrak{P}_1'(\omega) = \begin{cases} 0, & w_n \in S_2, \\ \mathfrak{P}_1(\omega|_{S_1}), & w_n \in S_1 \end{cases} \qquad \mathfrak{P}_2'(\omega) = \begin{cases} 0, & w_n \in S_1, \\ \mathfrak{P}_2(\omega|_{S_2}), & w_n \in S_2 \end{cases}$$

where by $\omega|_{S_i}$, $i = 1, 2$ we denote an S_i-subword of ω (see §1.1).

Definition 4.3.4 Let $\tau = max(\tau_1', \tau_2')$

Theorem 4.3.5 The stopping time τ defined above is a strong uniform time. Moreover, if stopping times τ_1, τ_2 are perfect and respect extremal elements $\widehat{g}_1 \in G_1$, $\widehat{g}_2 \in G_2$, then τ is perfect and respects an extremal element $(\widehat{g}_1, \widehat{g}_2) \in G_1 \times G_2$.

Proof Observe that $(G_1, G_2) = (\langle S_1 \rangle, \langle S_2 \rangle)$ is an exact decomposition. By definition, the stopping times τ_1', τ_2' are strongly independent with conditional probability distribution U_{G_1} and U_{G_2} (see §3.5). By the Theorem 3.5.9, $\tau = max(\tau_1', \tau_2')$ is strong uniform which proves the first part of the Theorem.

By Theorem 3.3.3 to prove the second part of the Theorem all we need to prove is that the stopping time τ respects the extremal element $(\widehat{g}_1, \widehat{g}_2)$. In the language of the stopping probability function $\mathfrak{P}$ of a strong uniform time τ this means that $\mathfrak{P}(\omega) = 1$ for all $\omega \in S^\star(\widehat{g}_1, \widehat{g}_2)$. On the other hand, by definition of the $max(\tau_1', \tau_2')$ we know that for any ω as above $\mathfrak{P}_1'(\omega|_{S_1}) = 1$, $\mathbf{A}_1(\omega|_{S_1}) = 1$, $\mathfrak{P}_2'(\omega|_{S_2}) = 1$, $\mathbf{A}_2(\omega|_{S_2}) = 1$. Therefore by definition of the stopping probability function $\mathfrak{P}$ we have $\mathfrak{P}(\omega) = 1$ and we proved the second part of the Theorem. □

Theorem 4.3.5 has an obvious generalization for $(p_1, \ldots, p_n)$-weighted direct products of the walks. Now we would like to use our result to give a bound for the total separation of the direct products of the random walks.

Theorem 4.3.6 Let $\mathcal{W}$ be a $(p_1, \ldots, p_n)$-weighted direct product of n random walks $\mathcal{W}_1, \ldots, \mathcal{W}_n$, where $p_1 + \cdots + p_n = 1$. Denote s, $s_1, \ldots, s_n$ total separation distances for these walks. Then

$$s \leq \frac{s_1}{p_1} + \cdots + \frac{s_n}{p_n}$$

Proof Observe that $\mathcal{W}$ can be obtained by a sequence of weighted products:

$$\mathcal{W} = \left(\ldots \left(\left(\mathcal{W}_1 \times_{(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2})} \mathcal{W}_2 \right) \times_{(\frac{p_1+p_2}{p_1+p_2+p_3}, \frac{p_3}{p_1+p_2+p_3})} \mathcal{W}_3 \right) \cdots \times_{(1-p_n, p_n)} \mathcal{W}_n \right)$$

Therefore it is enough to prove the result for weighted direct products of just two random walks.

Consider two random walks $\mathcal{W}_1$, $\mathcal{W}_2$ and their $(p, 1-p)$-weighted direct product $\mathcal{W}$. Let τ_1, τ_2 be the perfect times for the random walks $\mathcal{W}_1$, $\mathcal{W}_2$.

Compute $E_1 = E(\tau_1'$ the expected number of steps of $\mathcal{W}$ before $\mathbf{A}_1 = 1$. Observe that the random walk $\omega_t|_{S_1}$ is just a rescaled random walk $\mathcal{W}_1$ (see §4.2). By the Theorem 4.2.1 we get a bound on E_1:

$$E_1 \le E(\tau_1) \cdot \frac{1 - \mathbf{p}_1(e_1)}{1 - (p\,\mathbf{p}(e_1) + (1-p))} \le \frac{s_1}{p}$$

Analogously for the similarly defined E_2 we get $E_2 \le \frac{s_2}{1-p}$.

If we start our walk $\omega_t|_{S_2}$ on G_2 only *after* we have $\mathbf{A}_1 = 1$. We get an easy bound:

$$E(\tau) \le E_1 + E_2$$

Therefore

$$s \le E(\tau) \le E_1 + E_2 \le \frac{s_1}{p} + \frac{s_2}{1-p}$$

This finishes the proof. □

Example 4.3.6 As in the Example 4.3.1, consider a random walk on $\mathbb{Z}_4$ which is also a direct product of two random walk on $\mathbb{Z}_2$. Since the total separation is equal to 1 for the random walk on $\mathbb{Z}_2$, we can use the Theorem above to bound the total separation for the random walk on $\mathbb{Z}_4$:

$$s \le \frac{1}{\frac{1}{2}} + \frac{1}{\frac{1}{2}} = 4$$

Recall that $s = 3$ (see Example 2.2.13).

Note also that the perfect time defined in Example 3.2.5 is different from the one defined by Definition 4.3.4 in this case. Indeed, the latter perfect time with probability $\frac{1}{2}$ stops at 0 after two walk steps of staying at 0. On the other hand, the first perfect time never stops 0 if the walk never left it.

Example 4.3.7 Consider a random walk $\mathcal{W}$ on an n-cube $\mathbb{Z}_2^n$ defined in Example 2.5.2. This random walk can be presented as an n-th power of a random walk on $\mathbb{Z}_2$. Let us construct a perfect time for $\mathcal{W}$. We can think of our walk as of the nearest neighbor random walk on a Cayley graph Γ which is a 1-skeleton of an n-dimensional cube. Indeed, define the walk as follows:

Start at $e = (0, 0, \ldots, 0)$. Each walk step

- choose a random coordinate i;
- flip a fair coin;
- if heads, move in the i-th direction, if tails stay.

Now define a stopping time τ as follows

- each time we choose a direction i, mark it;
- stop when all directions are marked.

This construction is due to Andre Broder (see [D], §4.*B*). One can show that this stopping time is exactly the same as the one defined by Definition 4.3.4. This stopping time is perfect simply because it respects the extremal element $(1, \dots, 1)$ (cf. Example 3.3.4).

Let us compute the expected stopping time for this walk. By construction, it is equal to the expected time to mark all the coordinates. Therefore

$$s = E(\tau) = \frac{n}{1} + \frac{n}{2} + \cdots + \frac{n}{n} = n \cdot \mathfrak{h}_n = O(n\, log(n))$$

Indeed, the expected time to mark the first coordinate is exactly $\frac{n}{1}$; the expected time to mark the second coordinate after the first one is marked, is exactly $\frac{n}{2}$; etc. This proves the result.

Observe that this is a version of the classical *coupon collector's problem* (see e.g. [Feller], §1.2). There is an Euler formula for the asymptotic behavior of $\mathfrak{h}_n$:

$$\mathfrak{h}_n = \ln(n) + \gamma + \frac{1}{2\,n} + O\left(\frac{1}{n^2}\right)$$

where $\gamma \approx 0.5772156649$ is the Euler-Mascheroni constant (see e.g. [BE], §1.1; [WW], §12.1). We have

$$s = n \cdot \mathfrak{h}_n = n \ln(n) + \gamma\, n + \frac{1}{2} + O\left(\frac{1}{n}\right)$$

It is not hard to show bounds for the remainder

$$-\frac{1}{2} < O\left(\frac{1}{n}\right) < 0$$

Therefore

$$n \ln(n) + \gamma\, n < s < n \ln(n) + \gamma\, n + \frac{1}{2}$$

Compare the right bound with the bound we get from the Theorem 4.3.6:

$$s \leq n\, \frac{1}{\frac{1}{n}} = n^2$$

Observe that from the Algorithm one can actually compute the whole separation series.

Theorem 4.3.8 For he random walk $\mathcal{W} = (\mathbb{Z}_2^n, S, \mathbf{P})$ the total separation series is given by the following formula:

$$s(x) = \sum_{k=1}^{n-1} \frac{(-1)^{k+1}\binom{n}{k}}{1 - \frac{n-k}{n}\, x} + (-1)^{n+1} = \frac{n}{1 - \frac{n-1}{n}\, x} - \frac{\binom{n}{2}}{1 - \frac{n-2}{n}\, x} + \cdots + (-1)^n\, \frac{n}{1 - \frac{1}{n}\, x} + (-1)^{n+1}$$

Proof By the Theorem 3.2.7 we have $s_m = P(\tau > m)$ where τ is the perfect time defined above. It means that s_m is the probability of not marking all the

coordinates. By the inclusion-exclusion principle (see e.g. [Stan], §2.1; [GJ] §2.2.29) this probability is equal to probability of marking exactly $n-1$ coordinates, minus the probability of marking exactly $n-2$ coordinates, plus etc. This gives us

$$s_m = \binom{n}{1} \left(\frac{n-1}{n}\right)^m - \binom{n}{2} \left(\frac{n-2}{n}\right)^m + \dots$$

Summing over all m: $s(x) = 1 + s_1 + s_2 + \dots$ we get the formula above. □

Corollary 4.3.9 The total separation s and the radius ρ of the random walk $\mathcal{W}$ are given by the formulas:

$$s = n \cdot \mathfrak{h}_n, \qquad \rho = \frac{n}{n-1}$$

Proof Clear. □

Example 4.3.10 Let $G = \mathbb{Z}_2^n$ and $S = \{e, s_i = (0, \dots, 1_i, \dots, 0)\}$ as above. Consider a lazy directed random walk $\mathcal{W} = (\mathbb{Z}_2^n, S, \mathbf{P})$, where $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(s_i) = p_i > 0$, $p_1 + \dots + p_n = \frac{1}{2}$. Then $\mathcal{W}$ is $(2\,p_1, \dots, 2\,p_n)$-weighted product of n copies of a standard random walk on $\mathbb{Z}_2$. Let τ be a perfect time defined again by marking the coordinates (or, equivalently, as $max(\tau_1, \dots, \tau_n)$ where τ_i is a "stop after one step" stopping time on i-th copy of $\mathbb{Z}_2$). By analogy with the symmetric case we have:

Theorem 4.3.11 Let $s(x)$ be the separation series of the random walk $\mathcal{W}$. Then

$$s(x) = \sum_{J=\{j_1, j_2, \dots\} \subset [n], J \neq [n], \emptyset} \frac{1}{1 - 2\,(p_{j_1} + p_{j_2} + \dots)\,x} \; + \; (-1)^n$$

Proof Analogous to the proof of the Theorem 4.3.8. □

Corollary 4.3.12 The total separation $s = s(1)$ and the radius ρ of the random walk $\mathcal{W}$ are given by the formulas:

$$s \leq \frac{1}{2\,p}\,\mathfrak{h}_n, \qquad \rho = \frac{1}{1-2\,p}$$

where $p = \min_{i \in [n]} p_i$.

Proof Clear. □

Remark 4.3.13 A particular case of the above defined walk $\mathcal{W}$ with probabilities p_i as in Zipf's law, was considered in [Dc].

Example 4.3.14 Let S' be another generating set of an n-cube $S' = \{e, s_1, \dots, s_n\}$, where

$$\begin{aligned} s_1 &= (1_1, \dots, 1_m, 0_{m+1}, \dots, 0_n) \\ s_2 &= (0_1, 1_2, \dots, 1_{m+1}, 0_{m+2}, \dots, 0_n) \\ &\dots \\ s_n &= (1_1, \dots, 1_{m-1}, 0_m, \dots, 0_{n-1}, 1_n) \end{aligned}$$

and m and n are relatively prime.

Consider a random walk $\mathcal{W}' = (\mathbb{Z}_2^n, S', \mathbf{P}')$ $\mathbf{p}'(e) = \frac{1}{2}$. This walk is due to Persi Diaconis (see [DP]).

Observe that the walk $\mathcal{W}'$ is equivalent to the walk $\mathcal{W}$ from the previous example. Therefore the total separation $s' = s = n \cdot \mathfrak{h}_n$.

We will define now another type of random walk on $G = G_1 \times G_2$. Suppose G_2 is abelian. Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$. Denote by $S_2 = S \cap \big(\{e_2\} \times G_2\big)$, $S_1 = S \setminus S_2$, $S_1' = S_1|_{G_1}$ - the projection of the set S_1 on the first coordinate. Now suppose S_2 generates G_2. Then S_1' generates G_1 as well.

Let $p = \sum_{s \in S_1} \mathbf{p}(s)$, $\mathbf{P}_1 = \frac{\mathbf{P}|_{S_1}}{p}$, $\mathbf{P}_2 = \frac{\mathbf{P}|_{S_2}}{1-p}$. We can consider random walks $\mathcal{W}_1 = (G_1, S_1', \mathbf{P}_1)$, $\mathcal{W}_2 = (G_2, S_2, \mathbf{P}_2)$.

Theorem 4.3.15 Suppose G_2 is abelian. The total separation of the random walk $\mathcal{W} = (G_1 \times G_2, S, \mathbf{P})$ is bounded by

$$s \leq \frac{s_1}{p} + \frac{s_2}{1-p}$$

where s_1, s_2 are the total separations of the walks $\mathcal{W}_1$, $\mathcal{W}_2$.

Proof Denote by τ_1, τ_2 perfect times for the random walks $\mathcal{W}_1$, $\mathcal{W}_2$. Define time-invariant stopping times τ_1', τ_2' as above. These are stopping for the random walk $\mathcal{W}$.

Since G_2 is abelian, we have $s_1 \cdot s_2 = s_2 \cdot s_1$ for all $s_1 \in S_1$, $s_2 \in S_2$. Therefore τ_1' and τ_2' are strongly independent and the stopping time $\tau = max(\tau_1', \tau_2')$ is strong uniform. Therefore $s \leq E(\tau)$.

The rest of the proof (computation of the $E(\tau)$) is similar to the computation in the proof of the Theorem 4.3.6. $\square$

Theorem 4.3.15 gives an upper bound for the total separation s. We will show now that under certain conditions this upper bound is tight.

We say that an element $\widehat{g} \in G_1$ is *S-independent* if for every word

$$\omega = s_1' s_2' \dots s_l' \in (S_1')^\star(\widehat{g})$$

a corresponding element

$$\psi(g, \omega) = \gamma(s_1 s_2 \dots s_l)|_{G_2}$$

does not depend on ω.

If every element of G_1 is S-independent, this simply means that $\psi : G_1 \to G_2$ is a group homomorphism.

Theorem 4.3.16 Under the conditions of Theorem 4.3.15, let $\widehat{g}_1$, $\widehat{g}_2$ be the extremal elements of the walks $\mathcal{W}_1$, $\mathcal{W}_2$. Then $(\widehat{g}_1, \widehat{g}_2 + \psi(\widehat{g}_1))$ is an extremal element of the walk $\mathcal{W}$ and the strong uniform time τ defined in the proof of the Theorem 4.3.5 is perfect.

Proof The proof is analogous to the proof of the second part of the Theorem 4.3.5. Basically whenever the walk gets to $(\widehat{g}_1, \widehat{g}_2 + \psi(\widehat{g}_1))$ it means that the associated walk $\mathcal{W}_1$ gets to $\widehat{g}_1$ and the walk $\mathcal{W}_2$ gets to $\widehat{g}_2$. Therefore we must always stop there. $\square$

Example 4.3.17 Let $G = \mathbb{Z}_n \times \mathbb{Z}_n$. Fix $0 \le a < n$. Consider the generating set

$$S = \{(0,0),\ (0,\pm 1),\ (1,a),\ (-1,a^2)\}$$

and probability distribution $\mathbf{P}$:

$$\mathbf{p}(0,0) = \frac{1}{2},\ \mathbf{p}(0,\pm 1) = \mathbf{p}(1,a) = \mathbf{p}(-1,a^2) = \frac{1}{8}$$

Take $p = \frac{1}{2}$. Then $\mathcal{W}_1 \cong \mathcal{W}_2$ is a random walk on $\mathbb{Z}_n$ and we get

$$s \le \frac{2}{\frac{1}{2}} \left(\frac{n^2}{6} + \frac{1}{3} \right) = \frac{2}{3}\, n^2 + \frac{4}{3}$$

4.4 Semidirect product construction.

Let G_1, G_2 be two finite groups, $G = G_1 \ltimes_f G_2$, where $f : G_1 \to Aut(G_2)$. By $g_2^{g_1}$ we denote an element $f(g_1^{-1})[g_2]$. Then multiplication in G is given by

$$(a_1, a_2) \times (b_1, b_2) = (a_1\, b_1, a_2\, b_2^{a_1})$$

Denote $G_1' = \{(g_1, e_2), g_1 \in G_1\}$, $G_2' = \{(e_1, g_2), g_2 \in G_2\}$ (see §4.3). Given a set of generators S of the group G denote $S_1 = S \cap G_1'$, $S_2 = S \cap G_2'$. Denote $S_1' = S_1|_{G_1}$, $S_2'|_{G_2}$.

We say that a random walk $\mathcal{W} = (G, S, \mathbf{P})$ is G_1-*symmetric* if for all $s \in S$, $g \in G_1$ we have $s^g \in S$ and $\mathbf{p}(s^g) = \mathbf{p}(s)$.

Recall that two random walks $\mathcal{W}_1 = (G_1, S_1, \mathbf{P}_1)$ and $\mathcal{W}_1 = (G_2, S_2, \mathbf{P}_2)$ are called *equivalent* $\mathcal{W}_1 \simeq \mathcal{W}_2$) if there is a one-to-one map $\phi : G_1 \to G_2$ which maps probability distribution Q_1^k into Q_2^k for all $k > 0$.

Theorem 4.4.1 Let $\mathcal{W} = (G, S, \mathbf{P})$ be a G_1-symmetric directed random walk such that $S = S_1 \cup S_2$. Then $\mathcal{W}$ is equivalent to the weighted direct product of random walks

$$\mathcal{W} \simeq \mathcal{W}_1 \times_{(p,1-p)} \mathcal{W}_2$$

where $\mathcal{W}_1 = (G_1, S_1', \mathbf{P}_1)$, $\mathcal{W}_2 = (G_2, S_2', \mathbf{P}_2)$, $p = \sum_{s \in S_1} \mathbf{p}(s)$, $\mathbf{P}_1 = \frac{\mathbf{P}|_{S_1'}}{p}$, $\mathbf{P}_2 = \frac{\mathbf{P}|_{S_2'}}{1-p}$.

Proof Consider an obvious map $\phi : G \to G_1 \times G_2$ which maps $(g_1, g_2) \in G$ into $(g_1, g_2) \in G_1 \times G_2$. We claim that this map establishes the equivalence of the random walk $\mathcal{W}$ and a weighted direct product of the walks on G_1, G_2 with conditional probability distributions $\mathbf{P}_1$, $\mathbf{P}_2$.

By analogy with Example 2.1.3 let us look at the Caley graphs Γ and Γ' of groups G and $G_1 \times G_2$. If suffice to show that the probability $P(g \to h)$ of the walk $\mathcal{W}$ going from g to h, $g, h \in G$ is equal to the probability $P(\phi(g) \to \phi(h))$ of the walk $\mathcal{W}_1 \times_{(p,1-p)} \mathcal{W}_2$ going from $\phi(g)$ to $\phi(h)$.

There are two kinds of edges in Γ depending on whether a generator s is in S_1 or in S_2. If $s = (a_1, e_2) \in S_1$ we get an (oriented) edge

$$(g_1, g_2) \to (h_1, h_2) = (g_1, g_2) \cdot_G \ (a_1, e_2)$$

where the subscript G under multiplication indicates that the product is considered in group G. Since $(g_1, g_2) \cdot_G \ (a_1, e_2) = (g_1 a_1, g_2)$ it implies that $\phi(h) = \phi(s) \cdot_{G_1 \times G_2} \phi(g)$ i.e. we get an edge in Γ' which corresponds to a generator $\phi(s) \in S_1'$. It implies that both probabilities of going along the edge in Γ and along the corresponding edge in Γ' are equal to $\mathbf{p}(s) = p \cdot \mathbf{p}_1(\phi(s))$.

Consider now edges associated with generators $s = (e_1, a_2) \in S_2$:

$$(g_1, g_2) \to (h_1, h_2) = (g_1, g_2) \cdot_G \ (e_1, a_2)$$

We have $(h_1, h_2) = (g_1, g_2) \cdot_G \ (e_1, a_2) = (g_1, g_2 a_2^{g_1})$. Since $\mathcal{W}$ is G_1-symmetric, there exist a generator $s' = (e_1, a_2^{g_1}) \in S_2$. Thus Γ' contains an edge

$$(g_1, g_2) \to (h_1, h_2) = (g_1, g_2) \cdot_{G_1 \times G_2} \ (e_1, a_2^{g_1})$$

and again by G_1-symmetry the probabilities of going along the edges in Γ and Γ' are equal to $\mathbf{p}(s) = p \cdot \mathbf{p}_2(\phi(s'))$.

This proves the equivalence of the random walk $\mathcal{W}$ and a weighted direct product of the walks on G_1, G_2 with conditional probability distributions $\mathbf{P}_1$, $\mathbf{P}_2$. $\square$

Example 4.4.2 Let $G = DH_m$ be a dihedral group. Recall that $DH_m \simeq \mathbb{Z}_2 \ltimes \mathbb{Z}_m$ (see §1.2). Consider a symmetric set of generators

$$S = \{e = (0,0),\ (1,0),\ (0,1),\ (0,-1)\}$$

and a probability distribution

$$\mathbf{p}(e) = \frac{1}{2}, \qquad \mathbf{p}(1,0) = \mathbf{p}(0,\pm 1) = \frac{1}{6}$$

.

Observe that the random walk $\mathcal{W} = (G, S, \mathbf{P})$ is $\mathbb{Z}_2$-symmetric and $S = S_1 \cup S_2$. Using Theorem 4.4.1 we get

$$\mathcal{W} \simeq \mathcal{W}_1 \times_{(1/3,2/3)} \mathcal{W}_2$$

where $\mathcal{W}_1$ is a random walk on $\mathbb{Z}_2$, $\mathcal{W}_1$ is a random walk on $\mathbb{Z}_m$. This is easy to see also just by looking at the Cayley graph $\Gamma = \Gamma(G, S)$ which is an m-prism in both cases (see §1.2).

Observe that both walks $\mathcal{W}_1$, $\mathcal{W}_2$ have extreme elements. If m is even, by theorems 4.3.5, 4.1.7 and 2.2.11 we get the total separation s for the walk $\mathcal{W}$

$$s = \frac{s_1}{\frac{1}{3}} + \frac{s_2}{\frac{2}{3}} = 3 + \frac{3}{2}\left(\frac{1}{6}\, m^2 + \frac{1}{3}\right) = \frac{1}{4}\, m^2 + 3\frac{1}{2}$$

where s_1, s_2 are total separations for the random walks $\mathcal{W}_1$, $\mathcal{W}_2$ respectively.

Analogously for m odd we get

$$s = 3 + \frac{3}{2}\left(\frac{1}{6}\,n^2 + \frac{1}{6}\right) = \frac{1}{4}\,m^2 + 3\frac{1}{4}$$

Let $G' = G_1 \times G_2$. Define an obvious map $\phi : G \to G'$. Let $S' = \phi(S)$, $\mathbf{P}' = \phi(\mathbf{P})$. By $\mathcal{W}'$ denote a random walk $(G', S', \mathbf{P}')$.

Theorem 4.4.3 Let $\mathcal{W} = (G, S, \mathbf{P})$ be G_1-symmetric directed random walk on $G = G_1 \ltimes G_2$. Then $\mathcal{W}$ is equivalent to a random walk $\mathcal{W}'$ on a direct product $G_1 \times G_2$.

Proof The proof is analogous to the proof of the Theorem 4.4.1. □

Example 4.4.4 Let $G = DH_m$, $S = \{e = (0,0),\ (1,\pm a),\ (0,\pm 1)\}$, where $a \in \mathbb{Z}_m$. Consider the following probability distribution $\mathbf{P}$ on S: $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(0,\pm 1) = \frac{1}{6}$, $\mathbf{p}(1,\pm a) = \frac{1}{12}$.

Denote by $\mathcal{W} = (G, S, \mathbf{P})$. The random walk $\mathcal{W}$ is $\mathbb{Z}_2$-symmetric and therefore is equivalent to the associated random walk $\mathcal{W}'$ on $\mathbb{Z}_2 \times_{(1/3,2/3)} \mathbb{Z}_m$. Take $S_1 = \{(0,0),\ (1,\pm a)\}$, $S_2 = \{(0,0),\ (0,\pm 1)\}$. Since G_2 is abelian we can use the Theorem 4.3.5 to get a bound for the total separation.

Example 4.4.5 Let $G = \mathbb{Z}_p \ltimes \mathbb{Z}_p^2$, p - prime, where $(x,y)^z = (x, y + z\,x)$, $x, y, z \in \mathbb{Z}_p$. Define a generating set

$$S = \{(0,0,0), (\pm 1,0,0), (0,\pm 1,a), 0 \le a \le p-1\}$$

and a probability distribution $\mathbf{p}(0,0,0) = \frac{1}{2}$, $\mathbf{p}(\pm 1,0,0) = \frac{1}{8}$, $\mathbf{p}(0,\pm 1,a) = \frac{1}{8\,p}$.

Define a random walk $\mathcal{W} = (G, S, \mathbf{P})$. Observe that $\mathcal{W}$ is G_1-invariant. Therefore $\mathcal{W} \simeq \mathcal{W}_1 \times \mathcal{W}_2$ where $\mathcal{W}_1 = (\mathbb{Z}_p, S_1, \mathbf{P}_1)$, $\mathcal{W}_1 = (\mathbb{Z}_p^2, S_2, \mathbf{P}_2)$ $S_1 = \{0, \pm 1\}$, $S_2 = \{(0,0),\ (\pm 1, a)\}$, $\mathbf{p}_1(0) = \mathbf{p}_2(0) = \frac{1}{2}$, $\mathbf{p}_1(\pm 1) = \frac{1}{4}$, $\mathbf{p}_2(\pm 1, a) = \frac{1}{4\,p}$

The total separations s_1, s_2 for the walks $\mathcal{W}_1$, $\mathcal{W}_2$ are given by the Theorem 4.1.7 (see §4.1). If $p > 2$ we get

$$s_1 = s_2 = \frac{1}{6}\,p^2 + \frac{1}{6}$$

Therefore the total separation s for the walk $\mathcal{W}$, $p > 2$ is bounded as

$$s \le 2\,(s_1 + s_2) = \frac{2}{3}\,p^2 + \frac{2}{3}$$

If $p = 2$ we get $s_1 = s_2 = 1$ and

$$s \le 2\,(s_1 + s_2) = 4$$

Note that since $G \simeq U(3, \mathbb{F}_p)$ we can interpret the above result as finding the total separation for the random walk generated by matrices

$$\begin{pmatrix} 1 & \pm 1 & * \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \pm 1 \\ 0 & 0 & 1 \end{pmatrix}$$

where by " $*$ " we mean any element $a \in \mathbb{Z}_p$.

4.5 Wreath product construction.

Let $G \subset S_n$ be a *permutation group* with elements represented as permutations, H - any finite group. Denote by R the wreath product of the groups G and H. In other words, $R = G \ltimes H^n$ a semidirect product, where G acts on H^n by permuting copies of H (see §1.1).

Suppose we are given two directed random walks $\mathcal{W}_G = (G, S_G, \mathbf{P}_G)$ and $\mathcal{W}_2 = (H, S_H, \mathbf{P}_H)$. There are at least two natural ways to construct a random walk on R. In this section we are going to study both of them and give some bounds on their total separation. Let s_G and s_H denote the total separation for the walks $\mathcal{W}_G$ and $\mathcal{W}_H$.

Denote by $(\sigma; h_1, \dots, h_n)$, $\sigma \in G$ the elements of the group R. By definition, an identity element of R can be written as

$$e_R = (e_G; e_H, \dots, e_H)$$

where $e_G = (1, 2, \dots, n) \in G$ is an identity element of G, $e_H \in H$ is an identity element of H. The multiplication in the group R is given by

$$(\sigma; h_1, \dots, h_n) \cdot (\sigma'; h'_1, \dots, h'_n) = (\sigma\,\sigma'; h_1\, h'_{\sigma(1)}, \dots, h_n\, h'_{\sigma(n)})$$

Define two sets $S_1 \supset S_2$ of generators of R as follows:

$$S_1 = \{(s; e_H, \dots, e_H), (e_G; e_H, \dots, e_H, s', e_H, \dots, e_H),\ s \in S_G, s' \in S_H\}$$
$$S_2 = \{(s; e_H, \dots, e_H), (e_G; s', e_H, e_H, \dots, e_H),\ s \in S_G, s' \in S_H\}$$

where s' in the fist case can be everywhere and in the second place only on the first place.

We assume that G is *transitive* i.e. the orbit of 1 consist of all the elements $1, \dots, n$. In this case it is easy to see that S_2 generates the whole group R.

For any p, $0 < p < 1$ define probability distributions $\mathbf{P}_1$, $\mathbf{P}_2$ as follows:

$$\mathbf{p}_1(s; e_H, \dots, e_H) = p\mathbf{p}_G(s), \qquad \mathbf{p}_1(e_G; e_H, \dots, s', \dots, e_H) = \frac{1-p}{n}\,\mathbf{p}_H(s')$$
$$\mathbf{p}_2(s; e_H, \dots, e_H) = p\mathbf{p}_G(s), \qquad \mathbf{p}_2(e_G; s', e_H, \dots, e_H) = (1-p)\,\mathbf{p}_H(s')$$

where $s \in S_1$, $s' \in S_2$.

Definition 4.5.1 The random walks $\mathcal{W}_1 = (R, S_1, \mathbf{P}_1)$ and $\mathcal{W}_1 = (R, S_2, \mathbf{P}_2)$ are called *p-weighted semidirect products of type I* and *of type II* respectively:

$$\mathcal{W}_1 = \mathcal{W}_G \ltimes_p^I \mathcal{W}_H, \qquad \mathcal{W}_2 = \mathcal{W}_G \ltimes_p^{II} \mathcal{W}_H$$

By analogy with direct product, we drop the part "p-weighted" when $p = \frac{1}{2}$ and write simply $\mathcal{W}_G \ltimes^I \mathcal{W}_H$ or $\mathcal{W}_G \ltimes^{II} \mathcal{W}_H$.

Theorem 4.5.2 Random walk $\mathcal{W}_1$ is equivalent to the $(p, 1-p)$-weighted direct product of the walks $\mathcal{W}_G$ and $\mathcal{W}_H^n$:

$$\mathcal{W}_1 \simeq \mathcal{W}_G \times_{(p,1-p)} \mathcal{W}_H^n$$

Proof Observe that the $\mathcal{W}_1$ is G-symmetric. By Theorem 4.4.1 we have the result. $\square$

Theorem 4.5.3 Let s be the total separation for the walk $\mathcal{W}_1 = (R, S_1, \mathbf{P}_1)$. Then

$$s \le \frac{s_G}{p} + \frac{n^2\, s_H}{1-p}$$

Proof By Theorem 4.5.2 and Theorem 4.3.6, we have

$$s \le \frac{s_G}{p} + n \cdot \frac{s_H}{\frac{1-p}{n}} = \frac{s_G}{p} + \frac{n^2\, s_H}{1-p}$$

which proves the result. $\square$

Example 4.5.4 Let $R = \mathbb{Z}_n \ltimes \mathbb{Z}_2^n$ be a wreath product of $G = \mathbb{Z}_n$ and $H = \mathbb{Z}_2$. Consider a semidirect product $\mathcal{W}_1 = \mathcal{W}_G \ltimes^I \mathcal{W}_H$ where $\mathcal{W}_G$, $\mathcal{W}_H$ are the standard random walks on $\mathbb{Z}_n$ and $\mathbb{Z}_2$ respectively.

By Theorem 4.5.2 we have $\mathcal{W}_1 \simeq \mathcal{W}_G \times \mathcal{W}_H^n$. The random walk on $\mathbb{Z}_n$ was studied in §4.1. The random walk $\mathcal{W}_H^n$ on the n-cube $\mathbb{Z}_2^n$ was studied in Example 4.3.7. Combining the results we get

$$\begin{aligned} s &\le 2\, \left(\frac{1}{6}\, n^2 + \frac{1}{3} + n \ln(n) + \gamma\, n + \frac{1}{2} \right) \\ &\le \frac{1}{3}\, n^2 + 2\, n \ln(n) + 2\, \gamma\, n + \frac{5}{3} \end{aligned}$$

Note that we can get an easy lower bound $s > \frac{1}{3}\, n^2$. Indeed, the reduced random walk on $G = \mathbb{Z}_n$ is a rescaled random walk $\mathcal{W}_G$. We have

$$s \ge \frac{s_G}{p} > \frac{1}{3}\, n^2$$

Example 4.5.5 Let $R = \mathbb{Z}_n \ltimes \mathbb{Z}_m^n$ be a wreath product of $G = \mathbb{Z}_n$ and $H = \mathbb{Z}_m$. Consider a semidirect product $\mathcal{W}_1 = \mathcal{W}_G \ltimes^I \mathcal{W}_H$. From Theorem 4.1.7 we get an easy upper bound for the total separation:

$$s \le 2\, \left(\frac{1}{6}\, n^2 + \frac{1}{3} + \frac{1}{6}\, n^2\, m^2 + \frac{1}{3}\, n^2 \right) = \frac{1}{3}\, n^2\, m^2 + n^2 + \frac{2}{3}$$

Example 4.5.6 Let $R = B_n = S_n \ltimes \mathbb{Z}_2^n$ be the hyperoctahedral group, $\mathcal{W}_1 = \mathcal{W}_G \ltimes^I \mathcal{W}_H$, where $\mathcal{W}_H$ is a standard random walks on $\mathbb{Z}_2$, and $\mathcal{W}_G$ is a random walk on S_n generated by all transpositions: $\mathcal{W}_G = (S_n, S, \mathbf{P})$, where $S = \{e, (i,j), 1 \le i < j \le n\}$, $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(i,j) = \frac{1}{n\,(n-1)}$.

It is known (see e.g. [Bou], §8.13.3) that the hyperoctahedral group is a Weyl group of the symplectic group $Sp(2\,n)$. It can be defined as a group of symmetries

of the corresponding root system $\{\pm e_i, \pm e_i \pm e_j\}$. We can think of $\mathcal{W}_1$ as of the random walk generated by all the reflections over the roots.

From the Example 4.3.7 we have $s_H = n\,\mathfrak{h}_n$. In §5.2 we prove that $s_G \leq 2\,n\,\mathfrak{h}_n$. From here we get

$$s \leq 2\,(2\,n\,\mathfrak{h}_n + n\,\mathfrak{h}_n) = 6\,n\,\mathfrak{h}_n$$

and finally

$$s < 6\,n\ln(n) + 6\,\gamma\,n + 3$$

Example 4.5.7 Let $R = B_n = S_n \ltimes \mathbb{Z}_2^n$ be the hyperoctahedral group (see §1.2). Consider $\mathcal{W}_1 = \mathcal{W}_G \ltimes^I \mathcal{W}_H$, where $\mathcal{W}_H$ is a standard random walks on $\mathbb{Z}_2$, and $\mathcal{W}_G$ is a random walk on S_n generated by "star transpositions": $\mathcal{W}_G = (S_n, S, \mathbf{P})$, where $S = \{e, (1,i), 1 < i \leq n\}$, $\mathbf{p}(e) = \mathbf{p}(1,i) = \frac{1}{n}$.

From the Example 4.3.7 we have $s_H = n\,\mathfrak{h}_n$. In §5.1 we will show that $s_G \leq 2\,n\,\mathfrak{h}_n$. We have

$$s \leq 2\,(2\,n\,\mathfrak{h}_n + n\,\mathfrak{h}_n) = 6\,n\,\mathfrak{h}_n$$

and finally

$$s < 6\,n\ln(n) + 6\,\gamma\,n + 3$$

Example 4.5.8 Let $R = B_n = S_n \ltimes \mathbb{Z}_2^n$ be the hyperoctahedral group, $\mathcal{W}_1 = \mathcal{W}_G \ltimes^I \mathcal{W}_H$, where $\mathcal{W}_H$ is the standard random walk on $\mathbb{Z}_2$, and $\mathcal{W}_G$ is a random walk on S_n generated by Coxeter transpositions (see §1.2) : $\mathcal{W}_G = (S_n, S, \mathbf{P})$, where $S = \{e, (i,i+1), 1 \leq i < n\}$, $\mathbf{p}(e) = \frac{1}{2}$, $\mathbf{p}(i,i+1) = \frac{1}{2\,(n-1)}$.

As before, $s_H = n\,\mathfrak{h}_n$. In §5.3 we prove that $s_G \leq \frac{1}{2}n^4$. From here we get

$$s \leq 2\,\left(\frac{1}{2}n^4 + n\,\mathfrak{h}_n\right) = n^4 + 2\,n\,\mathfrak{h}_n$$

Remark 4.5.9 Wreath products of the second type defined above are considered in joint work with Astashkevich (see [AP]).

5. Random walks on the symmetric group

5.1 The case of the star transpositions.

Let $G = S_n$ be the symmetric group, $S = \{s_1 = e, s_2 = (1,2), \ldots, s_n = (1,n)\}$ - the set of star transpositions, $\mathbf{P}$ - uniform probability distribution:

$$\mathbf{p}(s_1) = \mathbf{p}(s_2) = \cdots = \mathbf{p}(s_n) = \frac{1}{n}$$

Consider a directed random walk $\mathcal{W} = (G, S, \mathbf{P})$. In order to compute the total separation s of this random walk we present an explicit construction of a strong uniform time for this walk.

The idea is to mark some elements while walking. The convention is that once an elements is marked, it stays marked forever.

Algorithm 5.1.1 Mark the element n.

- Choose a random i, $1 \le i \le n$.
- If the element $\sigma(1)$ is unmarked and $\sigma(i)$ is marked, mark $\sigma(1)$;
- If the element $\sigma(1)$ is unmarked and $i = 1$, mark $\sigma(1)$;
- $\sigma \leftarrow (1,i) \cdot \sigma$;
- If all the elements are marked, stop. Else, return to the beginning.

It might be not obvious that the above algorithm defines a strong uniform time. The idea of a proof is the following. We prove by induction that if we have k marked elements, they are in random order with respect to each other, even conditioned on knowing which elements are marked, which places they occupy, and the time of the observation. So basically when we get $k = n$ we get a random permutation. Here is how this argument can be made precise.

We start with a formal definition of a stopping time τ. Let $I \subset [n]$. By $\sigma(I)$ denote the subset $\{\sigma(i) | i \in I\}$.

Algorithm 5.1.2 Set $I = J = \{n\}$, $k = 1$, $\sigma = e$;

- Do while $k < n$:
  - Choose a random $i \in [n]$;
  - If $1 \notin I$ and $i \in I \cup \{1\}$, Do:
    - $I \leftarrow I \cup \{1\}$, $J \leftarrow J \cup \{\sigma(1)\}$, $k \leftarrow k+1$;
  - $\sigma \leftarrow (1,i)\,\sigma$, $I \leftarrow \sigma(I)$;
- End do; Stop.

Note that we multiplied a transposition from the right. The reason for that is quite simple: we would like to think that when we apply (i,j) we exchange elements in places i and j. In the algorithm above the set J played role of the set of marked elements, I the set of their positions.

Theorem 5.1.3 The stopping time τ defined by Algorithm 5.1.2 is strong uniform.

Proof Let k, $I, J \subset [n]$, $|I| = |J| = k$ be as in the algorithm. If the walk is at a permutation $\sigma \in S_n$ we have $J = \sigma(I)$. Let $\pi_k : I \to J$ be a one-to-one correspondence $\pi_k(i) = \sigma(i)$, $i \in I$. By $\Pi(I, J)$ denote the set of all such correspondences. We claim that

- At each time, conditional on k, I, J, the correspondence $\pi_k : I \to J$ is uniform.

The claim is proved by induction. At the beginning of the algorithm we have $k = 1$, $I = J = \{n\}$ so the claim is obvious. Assume it is true for $k = m$. The claim remains true until a new element is marked i.e. until

- $1 \notin I$ and $i \in I \cup \{1\}$

where i is as in the Algorithm. In other words, the claim remains true until we choose to exchange the unmarked element $\sigma(1)$ in the first place with either of the elements $j \in J$ or let it stay. By definition, each of these possibilities has probability $\frac{1}{n}$. Therefore the new correspondence $\pi_{m+1} : I \cup \{1\} \to J \cup \{\sigma(1)\}$ is uniform, which proves the claim.

Now observe that when the algorithm stops, we have $k = n$, $I = J = [n]$. The claim implies that the stopping time τ defined by the algorithm is strong uniform. □

Example 5.1.4 Let $n = 3$, $G = S_3$, $e = (1, 2, 3)$. Consider in detail how the Algorithm works. First, it marks the last element 3: $(1, 2, 3_m)$. Then it keeps exchanging 1 and 2 untill it finally chooses to either stay and mark the first element or to mark the first element and exchange it with 3. When it does, we get some element a (which is either 1 or 2 depending on the parity of the time), and a pair of two other elements in random order: $(*_m, a, *_m)$. Now the algorithm either stays or exchanges two marked elements until it exchanges a with an element in the first place : $(a, *_m, *_m)$. Observe that two marked element are still in random order with respect to each other. The algorithm stops after the next step. At this step we either let a stay at the first place with probability $\frac{1}{3}$ or exchange it with one of the marked elements with equal probability. But then no matter what a is, we would still get every permutation with equal probability. This illustrates the proof of Theorem 5.1.3.

Now compute the total separation. It takes on average $\frac{3}{2}$ steps before we mark the second element, $\frac{3}{1}$ steps before we move the only unmarked element in the first place, and one more step before we mark the last element. Therefore the total separation is $s \leq E(\tau) = 5\frac{1}{2}$.

Theorem 5.1.5 Let s be the total separation of the random walk $\mathcal{W}$. We have

$$s \leq 2\, n\, \mathfrak{h}_n - n - 2$$

Proof First of all observe that $\tau = \tau_2 + \tau_3 + \cdots + \tau_n$, where the stopping times τ_l, $2 \leq l \leq n$ are defined as follows:

- walk till we mark the l-th element

By the Theorem 5.1.3 we have

$$s \leq E(\tau) = E(\tau_2) + \cdots + E(\tau_n)$$

Compute $E(\tau_k)$. Break τ_k into a sum of two stopping times:

- (τ'_k): walk till $1 \notin I$.
- (τ''_k): walk till $i \in I \cup \{1\}$.

By construction, when $k > 2$ we have

$$E(\tau_k) = E(\tau'_k) + E(\tau''_k) = \frac{n}{n-k+1} + \frac{n}{k}$$

Observe that at the beginning $1 \notin I = \{n\}$ and

$$E(\tau_2) = E(\tau''_2) = \frac{n}{2}$$

We conclude

$$E(\tau) = \left(\frac{n}{n-2} + \frac{n}{n-3} + \cdots + \frac{n}{1}\right) + \left(\frac{n}{2} + \frac{n}{3} + \cdots + \frac{n}{n}\right) < 2\,n\,\mathfrak{h}_n - n - 2$$

This proves the result. □

Theorem 5.1.6 Let s be the total separation of the random walk $\mathcal{W}$. Then for $n > 3$

$$s \geq \frac{1}{3}\,n\,\mathfrak{h}_n$$

Proof To get a lower bound on total separation, let us use Theorem 3.6.5. Take B be a set of all permutations $\sigma \in S_n$ with no fixed points (see Example 2.6.7). The hitting time ht_B is greater or equal to the expected time before we apply all the generators $(1, i)$, $2 \leq i \leq n$. By the coupon collector's problem this gives us

$$ht_B \geq \frac{n}{n-1}\,(n-1)\,\mathfrak{h}_{n-1} = n\,\mathfrak{h}_n - 1$$

From Theorem 3.6.5 and Example 2.6.7 for $n \geq 4$ we have

$$s \geq \frac{|B|}{|G|}\,ht_B > \frac{n! - 1}{e\,n!}\,(n\,\mathfrak{h}_n - 1) > \frac{1}{3}\,n\,\mathfrak{h}_n$$

which proves the result. □

Remark 5.1.7 In [FOW] the authors analyze this random walk using Fourier transform technique. They give asymptotic formalas for the expected hitting times of each element depending on the conjugacy class to which it belongs. Although the formula $s = O(n\,log(n))$ is not stated explicitly it follows easily from their analysis. Note that if the hitting time and average hitting time are known, one can use Theorem 3.6.1 to get a bound on the total separation. Unfortunately the asymptotic estimates for the hitting time and average hitting time given in [FOW] are too weak to give a reasonable bound.

Note also that in this case the diameter set D_S is the set of all involutions with zero or one fixed points at 1 depending on parity (see [P2]). However the hitting time maximizes on elements from the other conjugacy classes. Therefore it is impossible to use the second part of Theorem 3.6.1 to obtain an exact value of the total separation since by the Proposition 3.3.6 all extremal elements must belong to the diameter set.

5.2 The case of all transpositions.

Let $G = S_n$ be the symmetric group, $S = \{e, (i,j), 1 \le i < j \le n\}$ - the set of all transpositions, $\mathbf{P}$ - uniform probability distribution:

$$\mathbf{p}(e) = \frac{1}{2}, \qquad \mathbf{p}(i,j) = \frac{1}{n\,(n-1)}$$

Consider a directed random walk $\mathcal{W} = (G, S, \mathbf{P})$. We would like to present an explicit construction of a strong uniform time for this walk and then compute the total separation.

We can think of our random walk in the following way:

- Choose a random $i \in [n]$. Choose a random $j \in [n] \setminus \{i\}$. Flip a fair coin.
- If heads, $\sigma \leftarrow (i,j)\,\sigma$. Return to the beginning.

As in §5.1 we can mark the elements that are in random relative order. Here is how it can be done.

Algorithm 5.2.1 Mark the element n.

- Choose a random i, $i \in [n]$; choose a random j, $j \in [n] \setminus \{i\}$;
- Flip a fair coin. Let m be the number of marked elements.
- If heads, the element $\sigma(i)$ is unmarked and $\sigma(j)$ is marked, mark $\sigma(i)$;
- If heads, the element $\sigma(j)$ is unmarked and $\sigma(i)$ is marked, mark $\sigma(i)$;
- If heads, $\sigma \leftarrow (i,j) \cdot \sigma$;
- If tails, the element $\sigma(i)$ is unmarked and $\sigma(j)$ is marked, mark $\sigma(i)$ with probability $\frac{1}{m}$;
- If tails, the element $\sigma(j)$ is unmarked and $\sigma(i)$ is marked, mark $\sigma(j)$ with probability $\frac{1}{m}$;
- If all the elements are marked, stop. Else, return to the beginning.

Note that as in the Algorithm 5.1.1 whenever we exchange a marked and an unmarked element, we always mark the latter. The main difference between Algorithm 5.2.1 and Algorithm 5.1.1 is that when we stay and choose a marked and an unmarked element we mark the latter not always, but with a probability $\frac{1}{m}$, where m is the number of marked elements.

Consider informally what happens when $n = 3$. We start at $(1,2,3)$. First, we mark 3. Then we keep walking until either pair $(1,3)$ or pair $(2,3)$ is chosen. Say, it's $(2,3)$ and we get $(1, *_m, *_m)$ or in other words we get either $(1,2,3)$ or $(1,3,2)$ with equal probability. The number of marked elements $m = 2$ now. When we choose $(2,3)$ then, nothing new hapens. Suppose we choose either $(1,2)$ or $(1,3)$, which could happen with equal probability. We have four equally likely posibilities:

- We choose either $(1,2)$ or $(1,3)$, stay and do not mark 1.
- We choose either $(1,2)$ or $(1,3)$, stay and mark 1.
- We choose either $(1,2)$, exchange them and mark 1.
- We choose either $(1,3)$, exchange them and mark 1.

Indeed, each of the last two possibilities occurs with probability $\frac{1}{6}$. The probability of choosing $(1,2)$ (or $(1,3)$), staying and marking 1 is equal to $\frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$. Therefore conditioned we choose 1, all four possibilities are equally likely.

Note that we mark a new element 1 in only he last three cases which correspond to either staying or exchanging 1 and 2 or exchanging 1 and 3. This gives us a uniform permutation of S_3 condition all three elements are marked now.

Again one can prove that the above algorithm defines a strong uniform time. By analogy with the arguments in §5.1 we claim that during the course of the algorithm all arrangements of marked elements are equally likely. Here is how it can be done formally. Let $I \subset [n]$. Let $\sigma(I)$ denote the subset $\{\sigma(i) | i \in I\}$.

Algorithm 5.2.2 Set $I = J = \{n\}$, $m = 1$, $\sigma = e$;

- Do while $m < n$:
  - Choose a random $i \in [n]$; choose a random $j \in [n] \setminus \{i\}$; Flip a fair coin.
  - If heads, $i \notin I$ and $j \in I$: $I \leftarrow I \cup \{i\}$, $J \leftarrow J \cup \{\sigma(i)\}$, $m \leftarrow m+1$;
  - If heads, $j \notin I$ and $i \in I$: $I \leftarrow I \cup \{j\}$, $J \leftarrow J \cup \{\sigma(j)\}$, $m \leftarrow m+1$;
  - If tails, $i \notin I$ and $j \in I$ Do with probability $\frac{1}{m}$:
    - $I \leftarrow I \cup \{i\}$, $J \leftarrow J \cup \{\sigma(i)\}$, $m \leftarrow m+1$;
  - If tails, $j \notin I$ and $i \in I$ Do with probability $\frac{1}{m}$:
    - $I \leftarrow I \cup \{j\}$, $J \leftarrow J \cup \{\sigma(j)\}$, $m \leftarrow m+1$;
  - If heads, $\sigma \leftarrow (i,j)\,\sigma$, $I \leftarrow \sigma(I)$;
- End do; Stop.

Theorem 5.2.3 The stopping time τ defined by Algorithm 5.2.2 is strong uniform.

Proof The proof is analogous to the proof of Theorem 5.1.3. Let k be the number of marked elements, $I, J \subset [n]$, $|I| = |J| = k$ be as in the algorithm. If the walk is at a permutation $\sigma \in S_n$ we have $J = \sigma(I)$. Denote by $\pi_k : I \to J$ the one-to-one correspondence $\pi_k(i) = \sigma(i)$, $i \in I$. By $\Pi(I,J)$ denote the set of all such correspondences. We claim that

- At each time, conditional on k, I, J, the correspondence $\pi_k : I \to J$ is uniform.

The claim is proved by induction. At the beginning of the algorithm we have $k = 1$, $I = J = \{n\}$ so the claim is obvious. Assume it is true for $k = m$. The claim remains true until a new element is marked i.e. until either of the following conditions holds:

- we exchange a marked element and an unmarked element
- we stay and mark an unmarked element

Let us compute the probabilities of each of these possibilities. The probability of exchanging an unmarked element $\sigma(i)$ and a marked element $\sigma(j)$ is equal to $\frac{1}{m(n-m)}$. By the algorithm, the probability of staying and marking an unmarked element $\sigma(i)$ is equal to $\frac{1}{m} \cdot \frac{m}{m(n-m)} = \frac{1}{m(n-m)}$. Therefore the new correspondence $\pi_{m+1} : I \cup \{i\} \to J \cup \{\sigma(i)\}$ is uniform, which proves the claim.

Now observe that when the algorithm stops, we have $k = n$, $I = J = [n]$. The claim implies that the stopping time τ defined by the algorithm is strong uniform. □

Theorem 5.2.4 Let s be the total separation of the random walk $\mathcal{W}$. We have

$$s \leq 2\, n\, \mathfrak{h}_n$$

Proof In the notation of the proof of Theorem 5.1.5 we have

$$s \leq E(\tau) = E(\tau_1) + \cdots + E(\tau_{n-1})$$

By construction

$$E(\tau_m) = \frac{1}{\frac{m\,(n-m)}{n\,(n-1)} + \frac{1}{m}\,\frac{m\,(n-m)}{n\,(n-1)}} = \frac{n\,(n-1)}{(m+1)\,(n-m)}$$

We conclude

$$s \leq \sum_{m=1}^{n-1} \frac{n\,(n-1)}{(m+1)\,(n-m)} = \frac{n\,(n-1)}{n+1} \sum_{m=1}^{n} \frac{1}{m+1} + \frac{1}{n-m} \leq 2\, n\, \mathfrak{h}_n$$

This finishes the proof. □

Theorem 5.2.5 Let s be the total separation of the random walk $\mathcal{W}$. Then for $n > 3$

$$s \geq \frac{1}{6}\, n\, \mathfrak{h}_n$$

Proof As in the proof of the Theorem 5.1.6, take B to be a set of permutations with no fixed points. The hitting time ht_B is greater or equal to the expected time to apply at least once a transposition (i, j) for every i, $1 \leq i \leq n$. The latter time can be thought as a version of the coupon collectors problem when each time we are given two different coupons. Later on we shall prove that the the expected time in this version is

$$E = \frac{n\, \mathfrak{h}_n}{\frac{n}{n} + \frac{n}{n-1}} > \frac{1}{2}\,(n-1)\, \mathfrak{h}_n$$

Therefore for $n \geq 4$

$$s \geq \frac{|B|}{|S_n|} ht_B > \frac{n! - 1}{e\, n!}\, \frac{1}{2}\,(n-1)\, \mathfrak{h}_n > \frac{1}{6}\, n\, \mathfrak{h}_n$$

which proves the result. □

Remark 5.2.6 This walk was introduced and analyzed by Diaconis and Shahshahani in [DSh1]. Our strong uniform time is not perfect, altough the bounds we get bound total separation by up to a constant factor. The stopping time arguments were used by Broder and Matthews (see [Matt1], [D], §4*B*). Note that in contrast with our construction, both approaches are asymmetric (they distinguish between left and right elements in a permutation). In [Matt1] Matthews combined his and Broder's stopping times to get an asymptotically tight upper bound.

5.3 The case of adjacent transpositions.

Let $G = S_n$ be the symmetric group, $S = \{s_0 = e, s_i = (i, i+1), 1 \le i < n\}$ - the set of adjacent transpositions, $\mathbf{P}$ - uniform probability distribution:

$$\mathbf{p}(e) = \mathbf{p}(i, i+1) = \frac{1}{n}$$

We present a construction of a strong uniform time τ for the random walk $\mathcal{W} = (G, S, \mathbf{P})$. Here is the idea of τ. We will be coloring elements from $[n]$ in four colors: white, red, blue and green such that at the beginning all the elements are white, at the end all the elements are blue, and in the middle there are white, blue, green elements and at most one red element. Each time we choose the leftmost white element and color it red. There are only white and blue elements at that point. This red element will be moving as a slow random walk on a line according to some explicit rule. While moving, it will bump into white element which will turn green afterwards. Once there are no white elements left, we color the red element blue and all the green elements white again. We stop when all the elements are blue.

Algorithm 5.3.1 Color all elements white.

- While there are still white elements, Do:
  - Pick the leftmost white element. Color it red.
  - While there are still white elements, Do:
    - Suppose a red element is at place j.
    - Choose a random i, $0 \le i < n$. Flip a coin with probability of heads

$$P = \frac{1}{1 + \#\text{green elements}}$$

    - If heads, $i = 0$, and the element $\sigma(j+1)$ is white, Do:
      - color $\sigma(j)$ green and $\sigma(j+1)$ red.
    - If $i = j$ and $\sigma(i+1)$ is white, color it green.
    - If $i > 0$, exchange $\sigma(i)$ and $\sigma(i+1)$.
  - End Do.
  - Color red element blue. Color green elements white.
- End Do. Stop.

The reason why this algorithm defines a strong uniform time is the following. We use induction to prove that at each time the red element could be either of the white elements it came across with equal probability (those that didn't become red became green). When it turns blue it means that this is a random element among those that are not blue. This means that the first element a_1 which becomes blue is a random element $a_1 \in [n]$, the second blue element is a random element $a_2 \in [n] \setminus \{a_1\}$, etc. Note that the positions of the blue elements conceivably depend on time, but since they do not depend on the elements themselves, τ is strong uniform.

Here is a formal definition of τ and proof that it is strong uniform. We use numbers $0, 1, 2, 3$ to indicate colors white, red, blue and green respectively. The color of the element in place i is indicated by the number c_i. Let k indicate the number of blue elements. Let m indicate the number of green elements. Let j indicate the position of a red element.

Algorithm 5.3.2 Set $\sigma = e$, $c_1 = \cdots = c_n = 0$, $k = 0$.

- Do while $k < n$:
  - $j \leftarrow \min\{i \in [n], c_i = 0\}$, $c_j \leftarrow 1$, $m \leftarrow 0$;
  - Do while $m < n - k - 1$:
    - Choose a random i, $0 \le i < n$;
    - If $i = 0$, $c_{j+1} = 0$ Do with probability $\frac{1}{m+1}$:
      - $c_j \leftarrow 3$, $c_{j+1} \leftarrow 1$, $j \leftarrow j + 1$, $m \leftarrow m + 1$;
    - If $i = j$, $c_{j+1} = 0$ Do : $c_{j+1} \leftarrow 3$, $m \leftarrow m + 1$;
    - If $i > 0$ Do: $\sigma \leftarrow (i, i + 1)\,\sigma$, $c_i \leftrightarrow c_{i+1}$;
    - If $i > 0$, $i = j - 1$ Do: $j \leftarrow j - 1$;
    - If $i = j$ Do: $j \leftarrow j + 1$;
  - End Do.
  - $c_j \leftarrow 2$, $k \leftarrow k + 1$;
  - For $i = 1$ to n Do:
    - If $c_i = 3$, $c_i \leftarrow 0$;
- End Do. Stop.

Example 5.3.3 Let $n = 4$. We present an example of the working of the algorithm. Colors of the elements are denoted by indices to the permutation elements. Chosen transpositions are shown above the arrows along with the coin outcome (if appropriate).

$$(1,2,3,4) \xrightarrow{\text{coloring}} (1_r, 2_w, 3_w, 4_w) \xrightarrow{(2,3)} (1_r, 3_w, 2_w, 4_w) \xrightarrow{(1,2)} (3_g, 1_r, 2_w, 4_w)$$
$$\xrightarrow{e,\ \text{heads}} (3_g, 1_g, 2_r, 4_w) \xrightarrow{(2,3)} (3_g, 2_r, 1_g, 4_w) \xrightarrow{(3,4)} (3_g, 2_r, 4_w, 1_g) \xrightarrow{(2,3)} (3_g, 4_g, 2_r, 1_g)$$
$$\xrightarrow{\text{coloring}} (3_r, 4_w, 2_b, 1_w) \xrightarrow{e,\ \text{tails}} (3_r, 4_w, 2_b, 1_w) \xrightarrow{(3,4)} (3_r, 4_w, 1_w, 2_b) \xrightarrow{(1,2)} (4_g, 3_r, 1_w, 2_b)$$
$$\xrightarrow{e,\ \text{heads}} (4_g, 3_g, 1_r, 2_b) \xrightarrow{\text{coloring}} (4_r, 3_w, 1_b, 2_b) \xrightarrow{(2,3)} (4_r, 1_b, 3_w, 2_b) \xrightarrow{e} (4_r, 1_b, 3_w, 2_b)$$
$$\xrightarrow{(1,2)} (1_b, 4_r, 3_w, 2_b) \xrightarrow{(2,3)} (1_b, 3_g, 4_r, 2_b) \xrightarrow{\text{coloring}} (1_b, 3_r, 4_b, 2_b) \xrightarrow{\text{coloring}} (1_b, 3_b, 4_b, 2_b)\ \Diamond$$

Example 5.3.4 Let $n = 3$, $G = S_3$, $e = (1, 2, 3)$. Consider how Algorithm 5.3.1 works in this case. At the beginning we have $(1_r, 2_w, 3_w)$. We keep exchanging white elements 2 and 3 until the walk either stays or exchanges the red element in the first place (1) and an element in the second place. We get $(*_g, *_r, a_w)$, where a is either 2 or 3 depending on the parity of the number of steps made. Note that now the red element is either of the two non-white elements with equal probability.

We keep walking until we change colors next time. Observe that the white element cannot be to the left of the red element at that time since the only way it could get there is by exchanging with the red element which would lead to recoloring earlier. Now, a change of colors can happen in either of the two positions: $(*_g, *_r, a_w)$ and $(*_r, a_w, *_g)$. In each of the positions the probability of recoloring is $\frac{1}{2}$: we either exchange the red and the green element which happens with probability $\frac{1}{3}$, or we can stay and have a coin show heads, which happens with probability $\frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}$. Therefore, when recoloring in this case, we either stay with conditional probability $\frac{1}{3}$ or exchange white ang green elements with conditional probability $\frac{2}{3}$.

In case of the first position $(*_g, *_r, a_w)$, when recoloring, we always get $(*_g, b_g, *'_r)$ where $*'_r$ is either $*_r$ with probability $\frac{2}{3}$ or a_w with probability $\frac{1}{3}$. Therefore now the probability of the red element to be 1, 2 or 3 is equal to $\frac{1}{3}$. Analogously in the case of the second position $(*_r, a_w, *_g)$, when recoloring, we always get $(b_g, *'_r, *_g)$. Again, the probability of the red element to be 1, 2 or 3 is equal to $\frac{1}{3}$.

Note that since there are no white element left the algorithm changes color of the red element into blue and colors all the green element back to white again. Then we color the leftmost white element red. We get either $(c_r, d_w, *'_b)$ or $(c_r, *'_b, d_w)$, where c, d are any elements which could depend on a blue element $*'_b$ in a complicated way. The element $*'_b$ is still uniform in [3] in each of the two possibilities.

We keep walking by exchanging the blue element with either red or white elements until finally we reach the position when we have red and white next to each other (actually white will be always to the right of the red) and either stay or exchange them. It could happen in either of the two positions: $(c_r, d_w, *'_b)$ or $(*'_b, c_r, d_w)$. In both cases The element $*'_b$ is uniform in [3] and after we recolor the other two elements into red and green $(e_g, *''_r, *'_b)$ or $(*'_b, e_g, *''_r)$ we get a new red element $*''_b$ which is either c or d with equal probability, and e is the remaining element. This means that we have a random permutation $\sigma \in S_3$. The algorithm now simply recolors all the elements blue and stops.

We showed that Algorithm 5.3.1 indeed defines a strong uniform time τ for the random walk $\mathcal{W}$. Observe that even in such a small case as this, the exact computation of the $E(\tau)$ is somewhat cumbersome. Note also that in for $n = 3$ the random walk $\mathcal{W}$ is equivalent to a random walk on $\mathbb{Z}_6$ which was studied in §4.1.

Theorem 5.3.5 The stopping time τ defined by the Algorithm above is strong uniform.

Proof Suppose we know the whole sequence of "colors" c_i of element in place i, $1 \le i \le n$ at any time prior to the stopping of the algorithm. We will show that even conditioned to that information, the stopping state is still a random permutation.

First, observe that once an element became blue it stays blue till the end. This means that if we know the whole sequence of colors, we can reconstruct where the element which became blue first, which became blue second, etc. Analogously with green elements. Once they become green, they stay green until we get a new blue element.

Let b be the number of blue elements, j be the position of the unique red element. Denote by $(i) = (i_1, i_2, \ldots, i_b)$ the places of first blue element, the second blue element, etc. Let $I = \{i_1, \ldots, i_b\}$. By J denote the set of places of green and red elements. Obviously, $j \in J$ if the walk is at permutation $\sigma \in S_n$. Also $J \subset [n] \setminus I$. Denote $m = |J|$. Note that m in the algorithm is equal to the number of green elements i.e. is less than $|J|$ by one.

Let $\pi : I \to [n]$ be an injective map $\pi(i_c) = \sigma(i_c)$, $1 \le c \le b$. By $\Pi(I)$ denote the set of all such correspondences. We claim that

- At each time, conditional on b, (i), j, J, the map $\pi : I \to J$ is uniform.

Suppose now we know a set $\Sigma(J) = \{\sigma(j_c), j_c \in J\}$. We claim that

- At each time, conditional on b, (i), j, J, $\pi : I \to [n]$, $\Sigma(J)$, the element $\sigma(j)$ is uniform in $\Sigma(J)$.

We prove both claims by induction. At the beginning we have $b = 0$, $j = 1$, $J = \{1\}$, so the base of induction is obvious.

Suppose both claims are true up to a certain time. The first claim will remain true until we color a new blue element. But this could happen only with the red element when all the other elements are either blue or green. By the second claim, $\sigma(j)$ is uniform in $\Sigma(J) = [n] \setminus \{\pi(i_c), 1 \le c \le b\}$. Therefore a new map $\pi' : I \cup \{j\} \to [n]$, $\pi'(j) = \sigma(j)$ is uniform in $\Pi(I \cup \{j\})$, i.e. the first claim remains true. Note that when the number of blue elements b increases, all the green elements become white, so the second claim holds automatically.

Now prove the second claim. Suppose it is true up to a certain time. It will remain true until we get a new green element. This could happen in either of two possibilities:

- The red element $\sigma(j)$ is exchanged with a white element $\sigma(j+1)$ and the white element gets the green color.
- The red element $\sigma(j)$ gets the green color, the white element $\sigma(j+1)$ gets the red color, and the walk stays.

Suppose we have a white element in place $j+1$, i.e. immediately to the right of the red element. By the algorithm, the probability of the first event is $P_1 = \frac{1}{n}$ and the probability of the second event is $P_2 = \frac{1}{m\,n}$. Therefore conditional on getting a new green element, the probability of the new red element to be in $\sigma(J)$ is $\frac{m}{m+1}$. Also the probability of the new red element to be $\sigma(j+1)$ conditional on getting a new green element is equal to $\frac{1}{m+1}$. From here and inductive assumption we get

- the new red element is uniform in $\Sigma(J \cup \{j+1\})$

which finishes proof of the second claim.

Now observe that the algorithm stops only when all the elements are blue. The first claim then implies that the stopping time τ defined by the algorithm is strong uniform. This finishes the proof. □

Theorem 5.3.6 Let s be the total separation of the random walk $\mathcal{W}$. Then

$$\frac{1}{16}\,n^3 \le s \le \frac{1}{2}\,n^4$$

Proof First, the upper bound is proved. By the Theorem 5.3.5 we have $s \le E(\tau)$. Now we need to estimate $E(\tau)$. Observe that the expected time for he life of a red element is bounded by the hitting time of the first element in the last place. We know that this hitting time is given by

$$E = \frac{n}{2}\,n\,(n+1)$$

Indeed, we can think of the movements of the first element as of a random walk on line with probability of moving $\frac{1}{n}$ in each direction. We stop once we hit either n or $-n-1$. It gives the formula above.

Now, we need to wait till all the elements are blue. Note that the last element gets blue right away since there are no white element left. This means that

$$E(\tau) \leq (n-1)\,E = \frac{n}{2}\,(n-1)\,n\,(n+1) < \frac{1}{2}\,n^4$$

which proves the upper bound.

As for the lower bound, let $n = 2\,r$, $B = \{\sigma \in S_n, \sigma^{-1}(1) \geq r\}$. In other words, B is a set of permutations with element 1 placed in the second half. By the argument above we have

$$ht_B = \frac{n}{2}\,r\,(r+1)$$

We conclude

$$s \geq \frac{|B|}{|S_n|}\,ht_B = \frac{1}{2}\,\frac{n}{2}\,r\,(r+1) \geq \frac{1}{16}\,n^3$$

Identically the same argument works for $n = 2\,r+1$. $\square$

Remark 5.3.7 In the proof of Theorem 5.3.5 we implicitly use an algorithm for generating permutations which is due to Persi Diaconis (see [Dm]). The idea of Diaconis' algorithm is based on the following identity:

$$\overrightarrow{\prod_{k=n-1\ldots 1}}\ \overrightarrow{\prod_{i=1\ldots k}}\left(\frac{1}{i+1}\,e + \frac{i}{i+1}\,(i,i+1)\right) = \frac{1}{n!}\sum_{\sigma\in S_n}\sigma$$

Similar identities were obtained earlier by Jucys and Nazarov (see [Jucys], [Naz]) and in a less explicit context for any root system by Demazure (see [Dem]).

Use of comparison technique or coupling arguments allows to reduce the upper bound to for the total separation to $O(n^3\,log(n))$ (see [DSC], [A1]). We believe that this kind of bound should be possible to obtain by refining our construction.

5.4 The case of the k-cycles.

Fix $n, k \in \mathbb{N}$, $1 < k \leq n$. Let $G = S_n$, S be the set of all k-cycles, k - even. Obviously, $|S| = (k-1)!\binom{n}{k}$. Consider the uniform random walk $\mathcal{W} = (S_n, S, \mathbf{P})$ where $\mathbf{P}$ is given by

$$\mathbf{p}(e) = \frac{1}{2}, \qquad \mathbf{p}(s) = \frac{1}{2\,(k-1)!\,\binom{n}{k}},\ s \in \check{S}$$

Observe that when $k = 2$ we get the random walk on S_n generated by all transpositions (see 5.2).

Theorem 5.4.1 Let s be the total separation of the random walk $\mathcal{W}$. Then

$$s \geq \frac{2}{3}\,\frac{\mathfrak{h}_n}{\mathfrak{h}_n - \mathfrak{h}_{n-k}}$$

Proof Let B be the set of permutations with no fixed points. Then by Theorem 3.6.5, since $n \geq 2$ we have

$$s \geq \frac{|B|}{|G|}\, ht_B \geq s \geq \frac{1}{3}\, ht_B$$

We say that the walk *touches* $i \in [n]$ at step j if the k-subset I chosen at step j contains i. Let E be the expected time before we touch all the elements. Obviously $ht_B \leq E$ since in order to hit the element with no fixed points we need to touch all the elements.

As in §5.2 we can think of E as the expected time to collect all the n coupons if at each time we are given k different coupons. This problem imbeds in the usual coupon collector's problem. Indeed, consider an expected time E_l to collect l different coupons in the usual coupon collector's problem. We have

$$E_l = \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{n-l-1} = n\,(\mathfrak{h}_n - \mathfrak{h}_{n-l})$$

Therefore the expected time E' to get all the coupons if given k different coupons at once is bounded as

$$E' \geq \frac{E_n}{E_k} = \frac{\mathfrak{h}_n}{\mathfrak{h}_n - \mathfrak{h}_{n-l}}$$

Note also that since $\mathbf{p}(e) = \frac{1}{2}$ we have $E = 2\,E'$. Combining the results we have

$$s \geq \frac{1}{3}\, ht_B \geq \frac{2}{3}\, E \geq \frac{1}{3}\, \frac{\mathfrak{h}_n}{\mathfrak{h}_n - \mathfrak{h}_{n-l}}$$

which finishes the proof. □

Remark 5.4.2 This walk was studied by Lulov and Roichman (see [Lulov], [Roi]), who found a similar lower bound for the total variation distance. Unfortunately the upper bound given in [Roi] is quite far apart from the above lower bound when $k = o(n)$. Recently Lulov and the author (see [LP]) proved that when $k > \frac{n}{2}$ the lower bound meets the upper bound up to a constant factor. Note also that the diameter of S_n in terms of k-cycles can be asymptotically smaller than the mixing time (see [Vishne] for examples and references).

It is possible to define a strong uniform time for this walk. Unfortunately our construction is too cumbersome and inexplicit, so in view of the above mentioned results we decided not to present it here.

5.5 The case of weighted transpositions.

Let $G = S_n$, $S = \{e, (i,j), 1 \leq i < j \leq n$ be the set of all transpositions and $\mathbf{P}$ be a strictly positive probability distribution:

$$\mathbf{p}(e) = \frac{1}{2}, \qquad \mathbf{p}(i,j) = p_{i,j}$$

where $p_{i,j} > 0$, $\sum_{1 \leq i < j \leq n} p_{i,j} = \frac{1}{2}$.

Let $\mathcal{W} = (G, S, \mathbf{P})$ be a lazy directed random walk. It is possible to modify the strong uniform time defined in §5.2 to work in this case. The idea is again to make

a use of Lemma 5.1.4. This is achieved by changing the marking rules so that the probability of marking an unmarked i is always the same no matter what j was chosen (see §5.2).

Let $i \notin I$, $I \subset [n]$. Denote $p_{i,I} = \min_{j \in I} p_{i,j}$, $\bar{p}_{i,I} = \sum_{j \in I} p_{i,j}$. Denote also $p_{j,i} = p_{i,j}$, $1 \le i < j \le n$.

Algorithm 5.5.1 Set $I = J = \{n\}$, $m = 1$, $\sigma = e$;

- Do while $m < n$:
  - Sample (i,j) from probability distribution $2\,p_{i,j}$; Flip a fair coin.
  - If heads, $i \notin I$ and $j \in I$ Do with probability $\frac{p_{i,I}}{p_{i,j}}$:
    - $I \leftarrow I \cup \{i\}$, $J \leftarrow J \cup \{\sigma(i)\}$, $m \leftarrow m+1$;
  - If heads, $j \notin I$ and $i \in I$: Do with probability $\frac{p_{j,I}}{p_{j,i}}$:
    - $I \leftarrow I \cup \{j\}$, $J \leftarrow J \cup \{\sigma(j)\}$, $m \leftarrow m+1$;
  - If tails, $i \notin I$ and $j \in I$ Do with probability $\frac{p_{i,I}}{\bar{p}_{i,I}}$:
    - $I \leftarrow I \cup \{i\}$, $J \leftarrow J \cup \{\sigma(i)\}$, $m \leftarrow m+1$;
  - If tails, $j \notin I$ and $i \in I$ Do with probability $\frac{p_{j,I}}{\bar{p}_{j,I}}$:
    - $I \leftarrow I \cup \{j\}$, $J \leftarrow J \cup \{\sigma(j)\}$, $m \leftarrow m+1$;
  - If heads, $\sigma \leftarrow (i,j)\,\sigma$, $I \leftarrow \sigma(I)$;
- End do; Stop.

Theorem 5.5.2 Algorithm 5.5.1 defines a strong uniform time τ for the random walk $\mathcal{W}$.

Proof The proof is identical to the proof of Theorem 5.2.3. We prove by induction the same claim as in the proof of the Theorem 5.2.3. The base of induction is obvious. Observe that conditional on having sampled $s = (i,j)$, $i \notin I$, $j \in I$ the probability of marking i is $\frac{p_{i,I}}{p_{i,j}}$. Therefore conditional on having having sampled an unmarked element i the probability of marking i is $p_{i,I}$, i.e. independent of j. Check that this probability is also equal to the probability of staying and marking. Indeed, the latter probability is equal to $\frac{p_{i,I}}{\bar{p}_{i,I}}$ times the probability of chosing i and a marked element $j \in I$, i.e. equal to $\frac{p_{i,I}}{\bar{p}_{i,I}} \cdot \big(\sum_{j \in I} p_{i,j}\big) = p_{i,I}$. This proves the induction step and concludes the proof. $\square$

Theorem 5.5.3 The total separation s of the random walk $\mathcal{W}$ is bounded by

$$\max\{\frac{n-1}{n\,p^*}; \frac{n\,\mathfrak{h}_n}{6}\} \le s \le \frac{2\,\mathfrak{h}_n}{p\,(n-1)}$$

where $p = \min_{i,j} p_{i,j}$, $p^* = \min_i \bar{p}_{i,[n]\setminus\{i\}}$, $1 \le i < j \le n$.

Proof For the upper bound, compare the probability of marking for each $i \notin I$ in the Algorithm 5.2.2 and in the Algorithm 5.5.1. Clearly, the second probability is at least $p\,n\,(n-1)$ times the first probability. Therefore by the Theorem 5.5.2 the expected time of τ is at most $\frac{1}{p\,n\,(n-1)}\,2\,n\,\mathfrak{h}_n$. Since s is strong uniform, we immediately get the upper bound.

The lower bound consists of two different bounds. The second comes again from comparing the stopping time τ and the stopping time defined by the Algorithm 5.2.2. Indeed, we use a known fact (see [MO]) that the expected time to touch all elements $i \in [n]$ minimizes when the probability distribution $2\,p_{i,j}$ is uniform.

For the first lower bound, take B be a set of permutations with i not a fixed point, i.e. $B = \{\sigma \in S_n, \sigma(i) \neq i\}$, where i is such that $p^* = \bar{p}_{i,[n]\setminus\{i\}}$. Then the hitting time ht_B is simply $\frac{1}{p^*}$. We conclude

$$s \geq \frac{|B|}{|G|}\, ht_B = \frac{n-1}{n}\,\frac{1}{p^*}$$

This finishes the proof. $\square$

Example 5.5.4 Let $p_{i,j} = \frac{Z}{j-i}$, $1 \leq i < j \leq n$, where

$$Z^{-1} = \sum_{1 \leq i < j \leq n} \frac{1}{j-i}$$

Since $p = \frac{Z}{n-1}$ we have

$$Z^{-1} = \mathfrak{h}_{n-1} + \cdots + \mathfrak{h}_1 \sim n \ln(n)$$

$$s \leq \frac{2\,\mathfrak{h}_n}{p\,(n-1)} = 2\,\mathfrak{h}_n\, Z^{-1} \sim 2\,n \ln^2(n)$$

which seems to be the right upper bound as n tends to infinity.

Remark 5.5.5 Weighted transpositions with any pattern can be analyzed via the path comparisons of Diaconis and Saloff-Coste (see [DSC]). Recently Handjani and Jungreis (see [HJ]) have shown that the second eigevalue for many underlying graphs (including the complete graph and trees) occurs at the n-dimensional representation. This means it is easy to bound the total variation distance tv_k (see §2.4) using the coupon collector's problem and the usual upper bound lemma.

5.6 The case of the weighted star transpositions.

It is not hard to modify the strong uniform time in the case of star transpositions as well. Let $G = S_n$, $S = \{s_1 = e, s_2 = (1,2), \ldots, s_n = (1,n)\}$, $\mathbf{p}(s_i) = p_i$, $p_1 \geq p_i > 0$, $i \in [n]$, $p_1 + \cdots + p_n = 1$. Denote $p_I = \min_{i \in I} p_i$, $I \subset [n]$.

Algorithm 5.6.1 Set $I = J = \{n\}$, $k = 1$, $\sigma = e$;

- Do while $k < n$:
  - Choose a random $i \in [n]$;
  - If $1 \notin I$, $i \in I \cup \{1\}$, Do with probability $\frac{p_I}{p_i}$:
    - $I \leftarrow I \cup \{1\}$, $J \leftarrow J \cup \{\sigma(1)\}$, $k \leftarrow k+1$;
  - $\sigma \leftarrow (1,i)\,\sigma$, $I \leftarrow \sigma(I)$;
- End do; Stop.

Theorem 5.6.2 The algorithm 5.6.1 defines a strong uniform time τ for the random walk $\mathcal{W} = (S_n, S, \mathbf{P})$.

Proof Analogous to the proof of the Theorems 5.1.3, 5.5.2. Each time whenever $1 \notin I$, we mark $\sigma(1)$ and exchange it with an element $\sigma(i)$, $i \in I$ with probability p_I. Since $p_1 \geq p_I$, we let the element $\sigma(1)$ stay and mark it with probability $p_1 \frac{p_I}{p_1} = p_I$. These arguments combined with the arguments in the proof of Theorem 5.1.3 prove the result. $\square$

Theorem 5.6.3 The total separation s of the random walk $\mathcal{W}$ is bounded by

$$\max\{\frac{1}{p}; \frac{n\,\mathfrak{h}_n}{3}\} \leq s \leq \frac{2\,\mathfrak{h}_n}{p^2\,n}$$

where $p = p_{[n]} = \min_{i \in [n]} p_i$.

Proof The proof of the lower bound is exactly the same as in proof of the Theorem 5.5.3. The proof of the upper bound is different because we can not mark a new element when there is a marked element in the first place. First we need to wait till there is an unmarked element there. However in contrast with the symmetric situation we can get a marked element in the first place without marking a new element.

Now suppose we have a set I of marked elements, $|I| = k$. The expected time to mark a new element E_k satisfies the following inequality:

$$E_k \leq \frac{1}{k\,p} + \frac{1}{(n-k-1)\,p} + ((k+1)\,p + (1-(k+1)\,p - (n-k-1)\,p)\ \cdot E_k)$$

Indeed, the expected time till the element in the first place is unmarked is at most $\frac{1}{k\,p}$ with the equality achieved when there is an $i \in I$ with $p_i = p$. Then, the expected time before we get the marked element in the first place is at most $\frac{1}{(n-k-1)\,p}$ (it is 0 when $k = n-1$). When we get the marked element on the first place we can either mark a new element or not. The probabilities of these events are at most $(k+1)\,p$ and $(1-(k+1)\,p-(n-k-1)\,p)$ respectively. In the latter case we need to start all over again. This gives the formula above.

From here we have

$$E_k(1-(1-n\,p)) \leq \frac{1}{k\,p} + \frac{1}{(n-k-1)\,p} + (k+1)\,p$$

$$E_k \leq \frac{1}{n\,p^2}\left(\frac{1}{k+1} + \frac{1}{n-k-1}\right) + \frac{k+1}{n}$$

Summing over k we get

$$s \leq E(\tau) = E_1 + \cdots + E_{n-1} \leq \frac{1}{n\,p^2}\left(2\,\mathfrak{h}_n - 1 - \frac{1}{n}\right) + \frac{(n+2)(n-1)}{2\,n}$$

Since $p \leq \frac{1}{n}$ we have

$$\frac{1}{n\,p^2}\left(1 + \frac{1}{n}\right) > \frac{(n+2)(n-1)}{2\,n}$$

This gives us

$$s \le \frac{2\,\mathfrak{h}_n}{n\,p^2}$$

which completes the proof of the Theorem. □

Example 5.6.4 Sometimes it is possible to improve the lower bound if more is assumed about the probability distribution $\mathbf{P}$. Let $\mathbf{p}(s_i) = p_i = \frac{Z_r}{i^r}$, where

$$Z_r^{-1} = \zeta_n(r) = \sum_{i=1}^{n} \frac{1}{i^r}$$

(cf. [Dc], §5). We have $p = \frac{Z_r}{n^r}$. The upper bound in Theorem 5.6.3 gives us

$$s \le \frac{2\,\mathfrak{h}_n}{p^2\,n} = 2\,\mathfrak{h}_n\,\zeta_n^2(r)\,n^{2\,r-1}$$

Suppose now $n = 2\,m$. Let B be the set of permutations with no fixed points at places $m+1, \dots, n$. The coupon collector's arguments (see Example 4.3.7) give us

$$ht_B \ge \frac{\zeta_n(r)\,n^r}{1\cdot 2^r} + \dots + \frac{\zeta_n(r)\,n^r}{\frac{n}{2}\cdot 2^r} = \frac{\zeta_n(r)\,n^r\,\mathfrak{h}_{(n/2)}}{2^r}$$

and by the Theorem 3.6.5 we have

$$s \ge \frac{|B|}{|G|}\,ht_B \ge \frac{\zeta_n(r)\,n^r\,\mathfrak{h}_{(n/2)}}{3\cdot 2^r}$$

When $r = 1$ we have $\zeta_n(1) = \mathfrak{h}_n$ and we get the lower and upper bounds

$$\frac{1}{6}\,n\,\mathfrak{h}_n\,(\mathfrak{h}_n - 1) \le s \le 2\,n\,\mathfrak{h}_n^3$$

which asymptotically gives us

$$C_1 \cdot n\ln^2(n) < s < C_2 \cdot n\ln^3(n)$$

When $r = 2$ we have $\zeta_n(2) < \frac{\pi^2}{6} = O(1)$. As n tends to infinity our bounds give us

$$O(n^2\ln(n)) < s < O(n^3\ln(n))$$

5.7 The case of the k-subsets of an n-set.

Let G be a finite group, $H \subset G$ a subgroup. Denote by C the set of right cosets G/H, i.e. every $c \in C$ has a form $c = g\,H$ for some $g \in G$. There is an obvious action of G on C: $g : g_1\,H \to (g\,g_1)\,H$, $g, g_1 \in G$.

Let $\mathcal{W} = (G, S, \mathbf{P})$ be a directed random walk. One can define a Markov chain on C as follows:

Start at $c_e = e\,H$ and move with transition probabilities

- $P(c_1 \to c_2) = \sum_{s \in S, s\, c_1 = c_2} \mathbf{p}(s)\,, \qquad c_1, c_2 \in C$

One can show that this Markov chain (denote it $\mathcal{W}/H = (G/H, S, \mathbf{P})$) has most of the properties the random walks have. In particular, it has a uniform stationary distribution and there is a similar notion of separation distance and total separation. We refer to papers [AD2, DF], and book [AF], §9 for the definitions and details about the strong stationary time approach in case of the general Markov chains. Some of the results we use that cannot be found in the references above can be easily deduced from our approach in this slightly more general but still very symmetric case.

To finish with the introduction, note that once we analyzed a random walk $\mathcal{W}$, we also have some knowledge of the Markov chain $\mathcal{W}/H$. In particular, if s_k (s'_k) is the separation distance of $\mathcal{W}$ ($\mathcal{W}'$ after k steps, then $s'_k \le s_k$. Analogously the strong uniform time for the random walk $\mathcal{W}$ can be projected to a strong uniform time for the Markov chain $\mathcal{W}/H$. Sometimes, however, one can improve these bounds by modifying the construction of the projected strong uniform time. here we present one of the examples when it can be done.

Let $G = S_n$, $H = S_k \times S_{n-k}$, $H \subset G$. We think of H as of a group of permutations that preserves the first k elements. Clearly $C = G/H$ is a set $\begin{bmatrix} n \\ k \end{bmatrix}$ of all k-subsets of $[n]$. Also $|C| = \left|\begin{bmatrix} n \\ k \end{bmatrix}\right| = \binom{n}{k}$, $c_e = [k]$.

Define a Markov chain $\mathcal{M}$ with the following transition probabilities:

$$P(I_1 \to I_2) = \begin{cases} \dfrac{1}{2}, & I_1 = I_2 \\ \dfrac{1}{2\,k\,(n-k)}, & |I_1 \cap I_2| = k-1 \\ 0, & \text{otherwise} \end{cases}$$

One can think of $\mathcal{M}$ as of a nearest neighbor random walk on a 1-skeleton of the hypersimplex (see e.g. [P]). We will show that $\mathcal{M}$ is exactly $\mathcal{W}/H$ where $\mathcal{W}$ is a rescaled random walk on S_n generated by all transpositions.

Algorithm 5.7.1 Start at $I = [k] \subset [n]$. Mark the elements in $[n] \setminus [k]$. Set $m = 0$.

- Choose a random $i \in I$, $j \in [n] \setminus I$. Flip a fair coin.
- If heads and the element i is unmarked, mark it.
- If tails and the element i is unmarked, mark it with probability $\frac{m+1}{n-k}$.
- If we mark i, $m \leftarrow m+1$.
- If heads, $I \leftarrow I \setminus \{i\} \cup \{j\}$.
- If all the elements are marked, stop. Else, return to the beginning.

One can prove that τ is strong uniform by the following argument.Let $J \subset [n]$ be the set of marked elements. We claim that at each time our k-subset I contains a random subset of $|J| - n + k$ marked elements even conditioned on knowing what are the elements J. The proof is similar to the analogous result for the random walk

on S_n generated by all transpositions. Before we define τ formally, let us show how the proof works in a small case. By m we denote the number of marked elements in our k-set I, i.e. $m = |J \cap I|$.

Example 5.7.2 Let $n = 4$, $k = 2$. Just for convenience we write the elements that are not in our 2-subset after the bar $|$. We start at $(1, 2 \,|\, 3_m, 4_m)$. After the first step we take either 1 or 2 and either exchange it with a random marked element 3 or 4 or leave it where it is. We always mark the element when we exchange it, and if we stay, mark it with probability $\frac{1}{2}$. This means that after the first step we get to the position $(a, *_m \,|\, *_m, *_m)$, where a is some element in $[2]$ and the other element in I is randomly chosen among the 3 remaining elements in $[4] \setminus \{a\}$.

Now we keep walking by taking the marked element and either leaving it in a 2-subset I or exchanging it with an element in $[4] \setminus I$, until we finally choose the remaining unmarked element a, which must be still in our 2-subset I. We always mark it then. The second element $*_m$ in I is a marked element which is still randomly chosen among the 3 remaining elements in $[4] \setminus \{a\}$. Let us compute the probabilities of each of the 2-subsets. Let for instance $a = 1$. Before we mark the last element we had $(1, 2 \,|\, 3, 4)$, $(1, 3 \,|\, 2, 4)$, $(1, 4 \,|\, 2, 3)$ with equal probability. Now we take 1 and either leave it in a 2-set I or exchange it with a random element in the complement $[4] \setminus I$. We get each of the $(1, 2 \,|\, 3, 4)$, $(1, 3 \,|\, 2, 4)$, $(1, 4 \,|\, 2, 3)$ with probability $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$. Also the probabilities of getting any of the three remaining 2-subsets $(2, 3 \,|\, 1, 4)$, $(3, 4 \,|\, 1, 2)$ and $(2, 4 \,|\, 1, 3)$ is equal to $2 \cdot \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{6}$. This proves that τ is indeed strong uniform in this case.

We now compute the expected number of steps for Algorithm 5.7.1 to work in this case. We need 1 step before we mark the first element, and the average of 2 steps before we take the unmarked element. We stop then. Thus we have the total separation is $s \leq E(\tau) = 1 + 2 = 3$.

Note that we can think of the Markov chain in this case as a nearest neighbor random walk on the octahedron. The furthest vertex of the octahedron from $(1, 2 \,|\, 3, 4)$ is $(3, 4 \,|\, 1, 2)$ which is an extremal element. One can see that our stopping time τ respects $(3, 4 \,|\, 1, 2)$ which implies that τ is perfect and $s = E(\tau) = 3$.

Algorithm 5.7.3 Set $I = [k]$, $J = [n] \setminus [k]$, $m = 0$.

- Do while $m < k$:
  - Choose a random $i \in I$, $j \in [n] \setminus I$; Flip a fair coin.
  - If heads, $i \notin J$ Do : $J \leftarrow J \cup \{i\}$; $m \leftarrow m + 1$.
  - If tails, $i \notin J$ Do with probability $\frac{m+1}{n-k}$: $J \leftarrow J \cup \{i\}$; $m \leftarrow m + 1$.
  - If heads, $I \leftarrow \big(I \setminus \{i\}\big) \cup \{j\}$.
- End do; Stop.

Theorem 5.7.4 The above algorithm defines a perfect time τ.

Proof The proof is similar to the proof of Theorem 5.2.3.

Let $m + n - k$ be the number of marked elements, J the set of marked elements, $|J| = n - k + m$. Suppose our chain is at k-subset I. Observe that by the algorithm, all unmarked elements must be in I, i.e. $I \setminus J \subset [k]$. This means that the m-subset

$R = J \cap I$ uniquely determines the set I. Let $\Pi(J, m)$ denote the set of all m-subsets R of a set J. We claim that

- At each time, conditional on m, J, the subset $R = J \cap I \subset J$ is uniform in $\Pi(J, m)$.

The claim is proved by induction. At the beginning of the algorithm we have $m = 0$, $I = [k]$, $J = [n] \setminus [k]$ so the claim is obvious. Assume it is true for $m = m'$. The claim remains true until a new element is marked i.e. until either of the following conditions holds:

- we exchange a marked element j and an unmarked element i
- we stay and mark an unmarked element i

Let us compute probabilities of each of these possibilities. The probability that the marked element j is exchanged with the unmarked element i is $P_1 = \frac{1}{2\,k\,(n-k)}$. The probability that the unmarked element i is marked and the chain stays is $P_2 = \frac{m'+1}{n-k}\,\frac{n-k}{2\,k\,(n-k)} = \frac{m'+1}{2\,k\,(n-k)}$. Therefore the probability that the unmarked element i is marked and remains in I conditioned that i is marked is $Q = \frac{m'+1}{n-k+m'+1}$.

We need to show that the probability Q is equal to $P(i \in R')$ where $R' \in \Pi(J \cup \{i\}, m'+1)$. This would imply that $R' = R \cup \{i\}$, $|R'| = m'+1$ is uniform in $\Pi(J \cup \{i\}, m'+1)$ and prove induction step. We have

$$P(i \in R') = \frac{\binom{n-k+m'}{m'}}{\binom{n-k+m'+1}{m'+1}} = \frac{m'+1}{n-k+m'+1} = Q$$

This proves the claim.

Now observe that when the algorithm stops, we have $m = k$, $J = [n]$. The claim implies that the stopping time τ defined by the algorithm is strong uniform.

Let us show that τ is perfect. Consider the element $\widehat{I} = \{k+1, k+2, \ldots, 2\,k\} \in C$, $\widehat{I} \subset [n] \setminus [k]$. Observe that in order for the chain to hit $\widehat{I}$ we need to touch all the elements in $[k]$ and therefore when we reach $\widehat{I}$ they are all marked. By construction of τ this means that when we hit $\widehat{I}$ we stop there. Therefore $\widehat{I}$ is an extremal element and by the Theorem 3.3.3 the strong uniform time τ is perfect. $\square$

From here we obtain the following result.

Theorem 5.7.5 Let s be the total separation for the Markov chain $\mathcal{M}$, $k \le n-k$. Then

$$s = \frac{2\,k\,(n-k)}{n+1}\,(\mathfrak{h}_n + \mathfrak{h}_k - \mathfrak{h}_{n-k})$$

Proof By Theorem 5.7.4 we have $s = E(\tau)$. By construction, $E(\tau) = E(\tau_1) + \cdots + E(\tau_k)$, where the stopping time τ_i, $1 \le i \le k$ is the stopping time when we mark the $(n-k+i)$-th element. From the algorithm we have:

$$E(\tau_i) = \frac{1}{\frac{k-i+1}{2\,k} + \frac{k-i+1}{2\,k}\,\frac{i}{n-k}} = \frac{2\,k\,(n-k)}{(k-i+1)(n-k+i)}$$

Thus we get

$$s = E(\tau) = E(\tau_1) + \cdots + E(\tau_k) = \sum_{i=1}^{k} \frac{2\,k\,(n-k)}{(k-i+1)(n-k+i)}$$
$$= \frac{2\,k\,(n-k)}{n+1} \sum_{i=1}^{k} \left(\frac{1}{k-i+1} + \frac{1}{n-k+i} \right) = \frac{2\,k\,(n-k)}{n+1} \left(\mathfrak{h}_n + \mathfrak{h}_k - \mathfrak{h}_{n-k} \right)$$

This proves the result. □

Corollary 5.7.6 Under the conditions above

$$k\,\mathfrak{h}_k < s < 4\,k\,\mathfrak{h}_k$$

Proof Clear. □

Remark 5.7.7 This Markov chain was introduced and analyzed by Diaconis and Shahshahani in [DSh2] (see also [D], §3F). They use there a Fourier transform technique and representation theoretic properties of Gelfand pairs. Their upper bound is slightly off when $k = o(n)$. This bound was subsequently improved by Greenhalgh in [Gr].

One can also generalize the observation made in Example 5.7.2 and show that the markoc chan $\mathcal{M}$ is equivalent to the nearest neighbor random walk on a *hypersimplex* $H_n^k \in \mathbb{R}^n$, $1 \le k \le n/2$ (see e.g. [EKK], §5.3), which can be defined by the following equations and inequalities:

$$\sum_{i=1}^{n} x_i = k \quad , \qquad 0 \le x_i \le 1, \quad 1 \le i \le n$$

5.8 The case of the semi-random transpositions.

Let $G = S_n$ be the symmetric group. Let $P_1, P_2, \ldots$ be an fixed infinite sequence of probability distributions on $[n]$. Consider the following stochastic process $\mathcal{M}$:

- At step k sample an element $j \in [n]$ from the distribution P_k.
- Choose a random i, $1 \le i \le n$.
- Apply the permutation (i, j). Return to the beginning.

By abuse of speech we call this stochastic process a *random walk generated by semi-random transpositions*. It is actually a random walk only if P_i, $i \ge 1$ are identical and independent of each other. Observe that if P_i, $i \ge 1$ are independent and uniform, we get a rescaled random walk generated by all transpositions. Also, when P_i is concentrated at 1 for all $i \ge 1$ we get the random walk generated by star transpositions.

We claim that it is possible to modify the construction of Algorithm 5.1.2 such that even in this generality it still defines a strong uniform time. In other words, we present a stopping rule that stops at a random permutation conditional the time of stopping.

Algorithm 5.8.1 Set $I = J = \{n\}$, $k = 1$, $m = 1$, $\sigma = e$;

- Do while $m < n$:
  - Choose a random $i \in [n]$;
  - Sample an element $j \in [n]$ from a probability distribution P_k;
  - If $j \notin I$ and $i \in I \cup \{j\}$, $I \leftarrow I \cup \{j\}$, $J \leftarrow J \cup \{\sigma(j)\}$, $k \leftarrow k+1$;
  - $\sigma \leftarrow (1,i)\,\sigma$, $I \leftarrow \sigma(I)$, $m \leftarrow m+1$;
- End do; Stop.

Theorem 5.8.2 The stopping time τ defined by the algorithm is strong uniform.

Proof The proof is identical to the proof of Theorem 5.2.3. Observe that in the proof of the claim in Theorem 5.2.3 we do not use any assumption on where the unmarked elements j come from. We were using only the fact that for each unmarked element j the probability of staying and marking it is equal to the probability of exchanging it with any marked element i and is equal to $\frac{1}{n}$. This is obviously true in the present algorithm. $\square$

Example 5.8.3 Consider a sequence of probability distributions $P_1, P_2, \ldots$ where P_k, $k \in \mathbb{N}$ is concentrated at a point $k \bmod n$. In notations of the Algorithm 5.8.1, each time we are forced to choose $j = k \bmod n$. This process $\mathcal{M}$ was defined in [AD1] and posed as an open problem.

Theorem 5.8.4 Let s be the total separation of the stochastic process $\mathcal{M}$. Then for any choice of P_i,

$$s \le 3\,n^2$$

Proof We use a generalization of Theorem 3.2.7: $s \le E(\tau)$ (see [AD2], [DF]). Now we need to compute the expected number of steps for the work of Algorithm 5.8.1 in this case.

Suppose we have m marked elements in $[n]$. Let E_m be the expected number of steps before we mark a new element. Let r_m be the expected time before the we get an unmarked element $j \in [n]$ in the worst case. Since the probability of marking is $\frac{m+1}{n}$, we immediately get

$$E_m \le r_m + \frac{m+1}{n} \cdot 1 + \frac{n-m-1}{n} \cdot E_m$$

and thus

$$E_m \le \frac{n}{m+1}\, r_m + 1$$

We now prove that $r_m \le 3\,m$. Indeed, suppose the current time is k and d is the smallest number such that the element at place $(k+d) \bmod n$ is unmarked. The probability that after d steps it is still unmarked is the probability that during these d steps it was not touched i.e. equal to $\left(1 - \frac{1}{n}\right)^d$. Since $d \le m \le n-1$ we get

$$r_m \le \frac{m}{\left(1 - \frac{1}{n}\right)^{n-1}} \le 3\,m$$

Therefore

$$E_m \leq 1 + 3\,n\frac{m}{m+1} < 3\,n$$

and finally

$$s \leq E_1 + \cdots + E_{n-1} \leq 3\,n\,(n-1)$$

which proves the result. □

Remark 5.8.5 Note that in case of the all transpositions the Algorithm 5.2.2 works faster than the Algorithm 5.8.1 above since in this case we are able to use all the symmetries of the walk. Namely, we mark a new element not only when j is unmarked and i is marked, but also when i is unmarked and j is marked.

We believe that for each n there is a constant $c(n)$ such that the expected time for the work of the Algorithm 5.8.1 is less than $c(n)$.

6. Geometric random walks

6.1 Full linear group.

In this section we present three different algorithm for generating random elements of the full linear group $G = GL(n;\mathbb{F}_q)$. By $M(n;\mathbb{F}_q)$ denote the set of all matrices over $\mathbb{F}_q$, $|M(n;\mathbb{F}_q)| = q^{n^2}$.

Trial and Error Algorithm 6.1.1

- Choose a random matrix $M \in M(n;\mathbb{F}_q)$; $d \leftarrow det(M)$.
- If $d \neq 0$, output M. Else, return to the beginning.

Theorem 6.1.2 Let $c(n;q)$ be the expected number of steps for the work of the algorithm. Then $c(n;q) \leq 2$ if $q \geq 3$, and $c(n;2) \leq 4$.

Proof The probability $p(n;q)$ of choosing a nonsingular matrix is equal to

$$\begin{aligned} p(n;q) &= \frac{|GL(n,\mathbb{F}_q)|}{q^{n^2}} = \frac{q^n-1}{q^n}\ \frac{q^n-q}{q^n}\ \dots\ \frac{q^n-q^{n-1}}{q^n} \\ &= \left(1-\frac{1}{q}\right)\left(1-\frac{1}{q^2}\right)\dots\left(1-\frac{1}{q^n}\right) \\ &> \prod_{i=1}^{\infty}\left(1-\frac{1}{q^i}\right) = 1-\frac{1}{q}-\frac{1}{q^2}+\frac{1}{q^5}+\frac{1}{q^7}-\dots \end{aligned}$$

The last product can be expanded as a sum using Euler's pentagonal theorem

$$\prod_{i=1}^{\infty}\left(1-\frac{1}{q^i}\right) = \sum_{m=0}^{\infty}\frac{(-1)^m}{q^{m(3m\pm1)/2}}$$

Therefore the desired probability can be bounded as

$$\begin{aligned} p(n;q) &> 1-\frac{1}{q}-\frac{1}{q^2}+\frac{1}{q^5}-\left(\frac{1}{q^6}+\frac{1}{q^7}+\frac{1}{q^8}+\dots\right) \\ &= 1-\frac{1}{q}-\frac{1}{q^2}+\frac{1}{q^5}-\frac{1}{q^6}\frac{1}{1-\frac{1}{q}} = 1-\frac{1}{q}-\frac{1}{q^2}+\frac{1}{q^5}\left(1-\frac{1}{q-1}\right) \end{aligned}$$

From here we get

$$p(n;q) > 1-\frac{1}{q}-\frac{1}{q^2}$$

Finally,

$$\begin{aligned} c(n;2) &< \frac{1}{p(n,2)} < \frac{1}{1-\frac{1}{2}-\frac{1}{4}} = 4 \\ c(n;q) &< \frac{1}{p(n,q)} < \frac{1}{1-\frac{1}{3}-\frac{1}{9}} = \frac{9}{5} < 2 \end{aligned}$$

for $q \geq 3$. This finishes the Proof. $\square$

Recall that by $l(\sigma)$, $\sigma \in S_n$ we denote the *length* of the permutation σ (see §1.2). One can also describe $l(\sigma)$ as follows:

$$l(\sigma) = |\{(i,j), 1 \leq i < j \leq n, \sigma(i) > \sigma(j)\}|$$

where each pair (i,j) in the set on the right hand side is called an *inversion*. Let $\mathbf{L}_z$ be the probability distribution on S_n:

$$\mathbf{L}_z(\sigma) = \frac{l(\sigma)}{(n!)_z}$$

where $z > 0$ and

$$(n!)_z = \sum_{\sigma in S_n} z^{l(\sigma)} = \prod_{i=1}^{n} \frac{z^i - 1}{z - 1}$$

Denote by $(\sigma \cdot Id)$ a matrix obtained from an identity matrix Id by permuting rows according to σ.

Bruhat Decomposition Algorithm 6.1.3 Let $z = \frac{1}{q}$.

- Sample random matrices $U \in U(n; \mathbb{F}_q)$, $B \in B(n; \mathbb{F}_q)$.
- Sample a permutation $\sigma \in S_n$ from the probability distribution $\mathbf{L}_z$.
- Output $M = U^T \cdot (\sigma \cdot Id) \cdot B$. End.

First, let us explain how we sample in the Algorithm. Sampling $U \in U(n; \mathbb{F}_q)$ and $B \in B(n; \mathbb{F}_q)$ from the uniform distribution is trivial (simply choose random entries of the matrices). Sampling a permutation σ from the probability distribution $\mathbf{L}_z$ is trickier. We start with sampling numbers $a_1, a_2, \ldots, a_n \in \mathbb{Z}_+$, $0 \leq a_i \leq i$, $P(a_i = j) = \frac{z^j}{1+z+\cdots+z^{i-1}}$. Sometimes sequences $(a) = (a_1, \ldots, a_n)$ are called *inversion vectors*. We present a bijection γ between sequences $(a) = (a_1, \ldots, a_n)$ and permutations $\sigma \in S_n$ such that $l(\gamma(a)) = a_1 + \cdots + a_n$. This gives us a direct procedure of sampling from S_n.

The bijection γ is constructed as follows. Put element 1 on a line. It has $a_1 = 0$ elements to the right of it. Put element 2 on a line to the left or to the right from 1 such that it has a_2 elements to the right of it, etc. Keep doing so till we get the desired permutation $\sigma = \gamma(a)$.

Now we are ready to formulate the following result.

Theorem 6.1.4 The Algorithm 6.1.3 outputs a random nonsingular matrix $M \in GL(n; \mathbb{F}_q)$.

This result was proved in [Ran] and [P]. For a general subgroup approach to generating of group elements see [DSh3]. Note that in the Bruhat Decomposition Algorithm we still need to multiply matrices which gives us an order of n^3 steps.

6.2 k-subspaces of $\mathbb{F}_q^n$.

Fix $k, n \in \mathbb{N}$, $k \le n-k$. Let $G = GL(n;\mathbb{F}_q)$, $H = GL(k;\mathbb{F}_q) \times GL(n-k;\mathbb{F}_q)$, $H \subset G$. We think of H as of a group of linear transformations that preserves some k-dimensional subspace and its orthogonal compliment. Clearly $Gr(n,k;\mathbb{F}_q) = G/H$ is a set of all k-dimensional subspaces of the vector space $\mathbb{F}_q^n$. Also $|Gr(n,k;\mathbb{F}_q)| = \binom{n}{k}_q$.

By analogy with §5.8, define Markov chain $\mathcal{M}$ on $Gr(n,k;\mathbb{F}_q)$ with the following transition probabilities:

$$P(V_1 \to V_2) = \begin{cases} \dfrac{1}{2}, & V_1 = V_2 \\ \dfrac{1}{2\,q\,(k)_q\,(n-k)_q}, & dim(V_1 \cap V_2) = k-1 \\ 0, & \text{otherwise} \end{cases}$$

This is a lazy random walk on $Gr(n,k;\mathbb{F}_q)$ which was introduced in [DSh2] and studied later in [Gr], [Ar]. We can think of this walk as follows:

- Start at $V \in Gr(n,k;\mathbb{F}_q)$.
- Choose a random $V' \in Gr(n,k;\mathbb{F}_q)$, $dim(V \cap V') = k-1$. Flip a fair coin.
- If tails, $V \leftarrow V'$. Return to the beginning.

The following algorithm defines a strong uniform time τ for the Markov chain $\mathcal{M}$.

Algorithm 6.2.1 Start at $V \in Gr(n,k;\mathbb{F}_q)$. Let $W = V$, $m = k$.

- Choose a random $V' \in Gr(n,k;\mathbb{F}_q)$, $dim(V \cap V') = k-1$.
- $W' \leftarrow W \cap V'$; $d \leftarrow dim(W')$. Flip a fair coin.
- If tails Do: $V \leftarrow V'$; $W \leftarrow W'$; $m \leftarrow d$.
- If heads and $d = m-1$, with probability $P(m,k,n) = \frac{(k-m+1)_q}{(n-m+1)_q}$ Do:
    - $W \leftarrow W'$; $m \leftarrow d$.
- If $m = 0$, stop. Else, return to the beginning.

Theorem 6.2.2 The above algorithm defines a perfect time τ.

Proof In this proof we follow [P]. We claim that at each moment we get a random k-dimensional subspace V containing W. Formally, let

$$Gr(n,k;W;\mathbb{F}_q) = \{V \in Gr(n,k;\mathbb{F}_q), W \subset V\}$$

be the set of k-subspaces containing W. Then

$$P(V = V_0 \,|W = W_0, t = t_0) = \begin{cases} \dfrac{1}{|Gr(n,k;W_0;\mathbb{F}_q)|}, & V_0 \in Gr(n,k;W_0;\mathbb{F}_q) \\ 0, & \text{otherwise} \end{cases}$$

We stop when $m = dim(W) = 0$. The claim therefore implies that the stopping time τ is strong uniform. The claim is proved by induction on the number of steps.

At the beginning $V = W$, $m = dim(W) = k$ and the claim is obvious. Suppose now that the claim is true after t steps. At the next step there are two principally different possibilities. First, if $d = dim(W') = m$ i.e. if $W' = W$ the claim still holds by symmetry. Indeed, it means we either stayed or moved to another randomly chosen subspace inside the $Gr(n, k; W; \mathbb{F}_q)$ which proves the claim in this case.

Now, suppose $d = dim(W') = m - 1$. Then by symmetry V' is a randomly chosen subspace in $Gr(n, k; W'; \mathbb{F}_q) \setminus Gr(n, k; W; \mathbb{F}_q)$. By construction we always have $W \leftarrow W'$ if tails and have $W \leftarrow W'$ if heads with probability $P(m, k, n)$. Analogously, by symmetry, if heads we have W' is a random $(m-1)$-dimensional subspace of W, and V is a randomly chosen subspace in $Gr(n, k; W; \mathbb{F}_q)$. Therefore if the probability $P(m, k, n)$ is given by

$$P(m, k, n) = \frac{Gr(n, k; W; \mathbb{F}_q)}{Gr(n, k; W'; \mathbb{F}_q)}$$

we have proved the claim. Indeed, if we have $W \leftarrow W'$, we need to equate the probabilities of getting a vector space V in $Gr(n, k; W'; \mathbb{F}_q) \setminus Gr(n, k; W; \mathbb{F}_q)$ and in $Gr(n, k; W; \mathbb{F}_q)$. Since we flip a fair coin, we immediately get $P(m, k, n)$ as above.

Let us compute $P(m, k, n)$. Define $FL(m, k, n, \mathbb{F}_q)$ as the set of sequences $\emptyset = V^0 \subset V^m \subset V^k \subset V^n = \mathbb{F}_q^n$, where $dim(V^i) = i$. It is easy to see that

$$|FL(m, k, n, \mathbb{F}_q)| = \binom{n}{k}_q \binom{k}{m}_q$$

since the number of ways to choose the sequence $V^m \subset V^k \subset V^n$ can be thought as the number of ways to choose $V^k \subset V^n$ times the number of ways to choose $V^m \subset V^k$.

We have a group $GL(n, \mathbb{F}_q)$ acting transitively on $FL(m, k, n, \mathbb{F}_q)$, which implies that the number $c(m, k, n; q) = |Gr(n, k; W; \mathbb{F}_q)|$ of ways to choose a k-dimensional subspace which contain a given m-space W is

$$c(m, k, n; q) = \frac{\binom{n}{k}_q \binom{k}{m}_q}{\binom{n}{m}_q} = \binom{n-m}{k-m}_q$$

From here we have

$$P(m, k, n) = \frac{c(m, k, n; q)}{c(m-1, k, n; q)} = \frac{\binom{n-m}{k-m}_q}{\binom{n-(m-1)}{k-(m-1)}_q} = \frac{(k-m+1)_q}{(n-m)_q}\,.$$

This proves the formula for $P(m, k, n)$ given in the algorithm.

Thus we have proved the claim, which consequently implies that τ is strong uniform. Now prove that τ is perfect. Pick any k-dimensional subspace V_1 such that $dim(V_0 \cap V_1) = 0$, where V_0 is the k-dimensional subspace where the chain starts. Such a space V_1 exists since $k \leq n - k$. We claim that once the walk gets to V_1, it stops there. Indeed, since W is a subspace of both V_0 and $V = V_1$, we have $W \subset V_0 \cap V_1 = \{\emptyset\}$. Thus $m = dim(W) = 0$ and the walk must stop once it gets to

V_1. Therefore τ respects V_1, and by Theorem 3.3.3 τ is perfect. This finishes the proof. $\square$

Theorem 6.2.3 Let s be the total separation of the Markov chain $\mathcal{M}$. Then

$$\frac{(k)_q}{q^k}\,\mathfrak{h}_n(q) \le s \le 2\,\frac{(k)_q}{q^k}\,\mathfrak{h}_n(q)$$

where

$$\mathfrak{h}_n(q) = \sum_{m=1}^{k} \frac{q^m}{(m)_q}$$

Proof By Theorem 6.2.2 the stopping time τ defined by the Algorithm 6.2.1 is perfect. Therefore $s = E(\tau)$. By construction,

$$E(\tau) = E(\tau_1) + E(\tau_2) + \cdots + E(\tau_k)$$

where τ_i, $1 \le i \le k$ is the expected time to have $m = dim(W)$ decrease from $k-i+1$ to $k-i$.

Let $m = dim(W)$, $1 \le m \le k$. The probability p_m of decreasing m at the next step is equal to

$$p_m = \frac{1}{2}\,(1 + P(m,k,n)) \cdot P(dim(W') = m-1)$$

Indeed, p_m is equal to the probability of tails and $dim(W') = m-1$ plus the $P(m,k,n)$ times the probability of heads and $dim(W') = m-1$. This gives the formula above. Therefore

$$\frac{1}{2}\,P(dim(W') = m-1) \le p_m \le P(dim(W') = m-1)$$

We next compute the probability $P(dim(W') = m)$. We have the number of ways to choose $V' \in Gr(n,k;\mathbb{F}_q)$ such that $dim(V \cap V') = k-1$, $V' \supset W$ is equal to the number of ways to choose a $(k-1)$-dimensional subspace U, such that $W \subset U \subset V$ times the number of ways to choose a k-dimensional subspace V' such that $V' \supset U$, $V' \ne V$. In the notation of the proof of Theorem 6.2.2, the first number is equal to $c(m,k-1,k;\mathbb{F}_q) = (k-m)_q$, and the second number is equal to $c(k-1,k,n;\mathbb{F}_q) - 1 = q\,(n-k)_q$. Therefore we have

$$P(dim(W') = m-1) = 1 - \frac{(k-m)_q\,q\,(n-k)_q}{q\,(k)_q\,(n-k)_q} = q^{k-m}\,\frac{(m)_q}{(k)_q}$$

and

$$\sum_{m=1}^{k} \frac{1}{P(dim(W') = m-1)} = \frac{(k)_q}{q^k} \sum_{i=1}^{m} \frac{q^m}{(m)_q} = \frac{(k)_q}{q^k}\,\mathfrak{h}_n(q)$$

We have

$$s = E(\tau) = \frac{1}{p_m} + \frac{1}{p_{m-1}} + \cdots + \frac{1}{p_1}$$

and from

$$\frac{1}{P(dim(W')=m-1)} \leq \frac{1}{p_m} \leq \frac{2}{P(dim(W')=m-1)}$$

we get the result. □

Corollary 6.2.4 For any $n \geq 2\,k$ the total separation satisfies $k \leq s \leq 4\,k$.

Proof The lower bound comes from the diameter of being k and therefore trivial. For the upper bound, from Theorem 6.2.3 we have

$$s \leq 2\,\frac{(k)_q}{q^k}\,\mathfrak{h}_n(q)$$

For any $m > 0$ we have $(m)_q = \frac{q^m-1}{q-1}$. From here we get

$$s \leq 2\,\frac{q^k-1}{q^k\,(q-1)} \sum_{m=1}^{k} \frac{q^m\,(q-1)}{q^m-1} = 2\,\left(1-\frac{1}{q^k}\right) \sum_{m=1}^{k} \frac{1}{1-\frac{1}{q^m}}$$

Since $q \geq 2$ we conclude

$$s \leq 2 \cdot 1 \cdot \sum_{m=1}^{k} \frac{1}{1-\frac{1}{2}} = 4\,k$$

which proves the upper bound. □

Remark 6.2.5 Using a different technique the similar bounds in this case were obtained by D'Aristotile in [Ar]. His arguments involve a Fourier transform approach by use of the properties of the Gelfand pairs which were earlier applied by Greenhalgh in this case (see [Gr]).

Random walk on k-subspaces is an example of the walk on distance regular graphs. For general results in this case see [Bel], [Chung].

Observe that when $q = 1$ we get the same bounds as in §5.7. In some sense, the Markov chain on k-subspaces of an n-space is indeed a q-analog of the Markov chain on k-subsets of an n-set. Analogously the perfect time constructed in here is a q-analoq of the perfect time constructed in §5.7. See [P] for more regarding this connection.

6.3 The case of the upper triangular matrices over $\mathbb{F}_q$ and its generalizations.

Let $G = U(n;\mathbb{F}_q)$ be the group of the upper triangular matrices over a finite field. Recall that $R(i,j;a)$, $1 \leq i < j \leq n$, $a \in \mathbb{F}_q$ denotes an upper triangular matrix with ones on the diagonal, a in place (i,j) and zeros elsewhere. Let

$$S = \{R(i,j;a), 1 \leq i < j \leq n, a \in \mathbb{F}_q\}$$

$$\mathbf{p}(e) = \frac{1}{q}, \quad \mathbf{p}(R(i,j;a)) = \frac{1}{q\binom{n}{2}},\ a \neq 0$$

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$ (see Example 2.5.8). We can think of this walk as follows

- Choose a random pair (i, j), $1 \le i < j \le n$
- Choose a random element $a \in \mathbb{F}_q$.
- Apply the generator $R(i, j; a)$ and return to the beginning.

Here is an idea of the strong uniform time for this walk. Denote by $I = \{(i, j), 1 \le i < j \le n\}$. Whenever we use a generator $R(i, j; a)$, mark a pair (i, j). Stop when all the pairs are marked.

One can formalize this idea in the following Algorithm. By J we denote the set of marked pairs, $m = |J|$, M - our upper triangular matrix.

Algorithm 6.3.1 Set $M = e$, $J = \emptyset$, $m = 0$.

- Choose a random pair $(i, j) \in I$. Choose a random $a \in \mathbb{F}_q$.
- If $(i, j) \notin J$ Do : $J \leftarrow J \cup \{(i, j)\}$; $m \leftarrow m + 1$.
- $M \leftarrow R(i, j; a) \cdot M$
- If $m = \binom{n}{2}$, stop. Else, return to the beginning of the Algorithm.

Theorem 6.3.2 The above Algorithm defines a strong uniform time τ.

Proof Suppose at step l we choose a pair (i_l, j_l) and a finite field element a_l. We need to prove that

$$P(M = \bar{M} | \tau = k) = \frac{1}{q^{\binom{n}{2}}}$$

for any $\bar{M} \in U(n; \mathbb{F}_q)$ We claim that a stronger property holds. Namely,

$$P(M = \bar{M} | i_1 = \bar{i}_1, j_1 = \bar{j}_1, \ldots, i_k = \bar{i}_k, j_k = \bar{j}_k) = \frac{1}{q^{\binom{n}{2}}}$$

for any fixed sequence of pairs $(\bar{i}_1, \bar{j}_1)$, $\ldots$, $(\bar{i}_k, \bar{j}_k)$ such that

$$\bigcup_{l=1}^{k} \{(\bar{i}_l, \bar{j}_l)\} = I$$

In other words, we claim that any sequence of sets $\{R(i_l, j_l, a), a \in \mathbb{F}_q\}$ such that $\cup_{l=1}^{k} \{(i_l, j_l)\} = I$, defines a uniform decomposition (see §1.1). Denote

$$M = M(a_1, \ldots, a_k) = R(i_1, j_1, a_1) \cdot \ldots \cdot R(i_k, j_k, a_k)$$

Define a linear order "$\prec$" on I as follows:

$$(i, j) \prec (i', j') \text{ if } j' - i' > j - i \text{ or } j' - i' = j - i, \ i < i'$$

Geometrically this means that the smallest element is in the upper left corner of the matrix. The elements increase along the first diagonal, then along the second diagonal, etc.

Observe that the matrix elements of $M = (m_{i,j})$ can be written in the form $m_{i,j} = a_{i,j} + f_{i,j}$, where $a_{i,j}$ is the sum of the a_l for all $1 \le l \le k$ such that $(i_l, j_l) = (i, j)$, and where $f_{i,j}$ is a function of the elements a_l such that $(i_l, j_l) \prec (i, j)$ which are independent of $a_{i,j}$. We have

$$\begin{aligned} P(M = \bar{M} \,|\, i_1 = \bar{i}_1, \ldots, j_k = \bar{j}_k) &= P(m_{1,2} = \bar{m}_{1,2} \,|\, i_1 = \bar{i}_1, \ldots, j_k = \bar{j}_k f_{1,2} = \bar{f}_{1,2}) \\ &\cdot P(m_{2,3} = \bar{m}_{2,3} \,|\, i_1 = \bar{i}_1, \ldots, j_k = \bar{j}_k, m_{1,2} = \bar{m}_{1,2}, f_{2,3} = \bar{f}_{2,3},) \cdot \ldots \\ &\cdot P(m_{1,n} = \bar{m}_{1,n} \,|\, i_1 = \bar{i}_1, \ldots, j_k = \bar{j}_k, m_{1,2} = \bar{m}_{1,2}, m_{2,3} = \bar{m}_{2,3}, \ldots, f_{1,n} = \bar{f}_{1,n}) \end{aligned}$$

where the product on the right hand side is taken in the order $\prec$ on I. Now, since the sequence (i_l, j_l), $1 \le l \le k$ contains each of the elements in I, all the probabilities on the right hand side are equal to $\frac{1}{q}$. This proves the claim which implies the Theorem. $\square$

Theorem 6.3.3 Let s be the total separation of the random walk $\mathcal{W} = (G, S, \mathbf{P})$. Then for any q

$$s \le \binom{n}{2} \mathfrak{h}_{\binom{n}{2}}$$

Proof From Theorem 6.3.2, we have $s \le E(\tau)$. For the $E(\tau)$, we again get a coupon collector's problem (see Example 4.3.7). This gives us the result. $\square$

It is not true that τ is a perfect time. One of the reasons is that the construction disregards the structure of the finite field. Consequently q does not appear in the formula $E(\tau)$. An advantage of this is that we can easily generalize the construction for a general type of Markov chains.

Definition 6.3.4 Let H be any finite group with the group operation described as addition. Fix $n \in \mathbb{N}$. Let $G = H^n$. Let $f_{i,j}$, $1 \le i < j \le n$ be any function of $a_1, \ldots, a_{i-1}$ which could also depend on time and additional random events. Consider a Markov chain $\mathcal{M}(G; (f_{i,j}))$ on H^n defined as follows:

- Start at $e_G = (e_H, \ldots, e_H)$.
- Choose a random $i \in [n]$. Choose a random $a \in H$.
- Let $a_i \leftarrow a_i + a$; $\quad a_j \leftarrow a_j + f_{i,j}(a_1, \ldots, a_{i-1}, t, a)$, $i < j \le n$.
- Return to the beginning.

Theorem 6.3.5 Let s be the total separation of the Markov chain $\mathcal{M}(G; (f_{i,j}))$. Then

$$s \le n\, \mathfrak{h}_n$$

Proof Indeed, consider the following stopping time τ:

- Mark $i \in [n]$ whenever we choose it and it's unmarked.
- Stop when all elements of $[n]$ are marked.

We claim that the stopping state is a random element of G even conditioned on knowing the sequence $(i_1, \ldots, i_k)$ of indices used before stopping. Indeed, for any

i, $1 \le i \le n$ there is a time l, $1 \le l \le k$ when we add a random element a to a_i. Therefore even conditioned on knowing the elements $a_1, a_2, \ldots, a_{i-1}$ at all times before the stopping, we still get a random element a_i at the end since we could add to a_i only elements that could depend on $a_1, a_2, \ldots, a_{i-1}$ but not on a. Thus the stopping element $(a_1, a_2, \ldots, a_n)$ has a random element a_1 conditioned on $\tau = k$, has a random element a_2 conditioned on a_1 and $\tau = k$, ... , has a random element a_n conditioned on $a_1, \ldots, a_{n-1}$ and $\tau = k$. This proves the claim. Therefore τ is indeed strong uniform.

We again use the coupon collector's argument to get

$$s \le E(\tau) = n\, \mathfrak{h}_n$$

which finishes the proof. $\square$

It turns out that in some cases the strong uniform time defined in the proof of the Theorem 6.3.5 is also perfect.

Example 6.3.6 Let $\mathbb{A} = \{c_1\, x + c_2\, x^2 + \cdots + c_m\, x^m\}$ be a set of polynomials of x with coefficients $c_i \in \mathbb{Z}_p$, with no free term and degree at most m, $x^{m+1} = 0$. Observe that $\mathbb{A}$ is a finite ring under addition and multiplication of polynomials. Consider a Markov chain $\mathcal{M} = (\mathbb{A}^n, (f_{i,j}))$ where the functions $f_{i,j} = f_{i,j}(a_1, \ldots, a_{i-1}, t, a)$ are of the form

$$f_{i,j} = a \cdot g_{i,j}(a_1, \ldots, a_{i-1}, t)$$

i.e. depend linearly on a.

We claim that in this case the total separation $s = n\, \mathfrak{h}_n$. Indeed, suppose the chain starts at $e = (0, \ldots, 0)$. Let us show that the element $g = (x, \ldots, x)$ is extremal. Indeed, for any (i, j) and $(a_1, \ldots, a_n)$ the element $f_{i,j}$ is a product of two polynomials and therefore has a degree either 0 or greater or equal to 1. Thus we must choose all the indices $i \in [n]$ before the chains gets to g. By construction the strong uniform time τ always stops then, which proves that g is an extremal element and that in this case τ is perfect.

We finish the section by considering another random walk on $G = U(n; \mathbb{F}_p)$, p-prime. Let

$$S = \{e, R(i, j; \pm 1), 1 \le i < j \le n\}$$
$$\mathbf{p}(e) = \frac{1}{2}, \quad \mathbf{p}(R(i, j; \pm 1)) = \frac{1}{4\binom{n}{2}}$$

Consider a random walk $\mathcal{W} = (G, S, \mathbf{P})$. We can think of this walk as follows:

- Choose a random pair (i, j), $1 \le i < j \le n$
- Flip a fair coin. Choose a random element $a \in \{1, -1\}$.
- If heads, apply the generator $R(i, j; a)$.
- Return to the beginning.

Theorem 6.3.7 Let s be the total separation of the random walk $\mathcal{W} = (G, S, \mathbf{P})$. Then

$$s \le \frac{1}{24}\,(p^2 + 2)\,(n^2 - n)^2$$

Proof Recall the notation in the proof of Theorem 6.3.2. Let I be the set of pairs (i,j), $1 \le i < j \le n$, "$\prec$" be the linear order on them.

Let $\tau_{i,j}$, $1 \le i < j \le n$ be a perfect time for the standard random walk on the copy of $\mathbb{Z}_p$ corresponding to the matrix element (i,j). Whenever we apply $R(i,j;\pm 1)$, we make a ± 1 step in the walk on the corresponding copy of $\mathbb{Z}_p$ (cf. §4.3). Define

$$\tau = \tau_{1,2} + \tau_{2,3} + \cdots + \tau_{1,n}$$

where the order on I is the linear order "$\prec$". We claim that τ is a strong uniform time.

First, observe that τ is a sum of time-invariant stopping times and therefore also time-invariant. We need to prove that τ is uniform. Indeed, when $\tau_{1,2}$ is stopped, the matrix element $m_{1,2}$ is uniform in $\mathbb{F}_p$. Therefore it will stay uniform in $\mathbb{F}_p$. After $\tau_{1,2} + \tau_{2,3}$ we get two random elements: $m_{1,2}$ and $m_{2,3}$, etc. Thus τ is uniform and therefore strong uniform.

Concluding the proof, we have:

$$s \le E(\tau) \le \frac{p^2+2}{6} \left(\frac{n^2-n}{2} \right)^2$$

which proves the result. □

Remark 6.3.8 Note that by analogy with §4.3 in the proof of the Theorem 6.3.7 we can take $\tau = max(\tau_{1,2}, \tau_{2,3}, \dots \tau_{1,n})$. When n is large compared to p this will improve bounds significantly. For examle, when $p = 2$ the last radom walk is equivalent to the random walk described in beginning of this section. The bound we get then is $s = O(n^2 \log(n))$ while Theorem 6.3.7 gives us $s = O(n^4)$.

As noted before (see Example 2.8.11), other types of walks on $U(n;\mathbb{F}_q)$ were studied by Diaconis and Saloff-Coste and by Stong (see [D-S-C], [Stong1]). Their techniques based on the bonds on eigenvalues is very different from ours.

6.4 The case of the upper triangular matrices over $\mathbb{F}_q$ as q grows.

Let $G = U(n;\mathbb{F}_q)$ be the group of the upper triangular matrices over the finite field. We have $|G| = q^{\binom{n}{2}}$. Recall that $R(i,j;a)$, $1 \le i < j \le n$, $a \in \mathbb{F}_q$ denotes an upper triangular matrix with ones on a diagonal, a on place (i,j) and zeros elsewhere. Let

$$S = \{R(i,i+1;a), 1 \le i < n, a \in \mathbb{F}_q\}$$

$$\mathbf{p}(e) = \frac{1}{q}, \quad \mathbf{p}(R(i,i+1;a)) = \frac{1}{q\,(n-1)},\ a \ne 0, 1 \le i < n$$

Theorem 6.4.1 Let s be the total separation of the random walk $\mathcal{W} = (G, S, \mathbf{P})$. Then

$$s \le \frac{n^4}{6} \cdot \left(1 + \frac{1}{q-1}\right)^{\binom{n}{3}}$$

Corollary 6.4.2 Under the conditions of Theorem 6.4.1, if $q > n^3$ then

$$s < \frac{1}{4}\, n^4$$

Proof Clear. □

The proof of Theorem 6.4.1 is based on the properties of a certain stopping time τ we define below.

Denote by τ_i, $1 \le i < n$ the following stopping times

- Walk till $R(i, i+1; a)$, $a \in \mathbb{F}_q$ is used. Stop.

Define

$$\mu_{j,l} = \tau_j + \tau_{j-1} + \cdots + \tau_l$$

where $1 \le l \le j < n$. Finally, define

$$\tau = (\mu_{1,1}) + (\mu_{2,1} + \mu_{2,2}) + \cdots + (\mu_{n-1,1} + \mu_{n-1,2} + \cdots + \mu_{n-1,n-1})$$

The stopping time τ is not strong uniform. It is not even time-invariant (see §3.5). We show that in a certain sense τ tends to strong uniform as q tends to infinity.

Lemma 6.4.3 Let τ be a stopping time defined above, ϱ - its stopping state. Then

$$P(\varrho = g | \tau = k) \ge \frac{1}{|G|} \cdot \left(1 - \frac{1}{q}\right)^{\binom{n}{3}}$$

for any $g \in G$, $k > 0$.

First, we deduce Theorem 6.4.1 from Lemma 6.4.3 and then prove the lemma.

Proof of Theorem 6.4.1 From Theorem 3.4.7 and Lemma 6.4.3 we have

$$s \le \frac{E(\tau)}{\left(1 - \frac{1}{q}\right)^{\binom{n}{3}}} = E(\tau) \cdot \left(1 + \frac{1}{q-1}\right)^{\binom{n}{3}}$$

By definition $r = E(\tau_i) = (n-1)$ for all $1 \le i < n$. Also $E(\mu_{j,l}) = (j - l + 1) \cdot r = (j - l + 1)\,(n-1)$. Finally

$$\begin{aligned} E(\tau) &= E(\mu_{1,1}) + \big(E(\mu_{2,1}) + E(\mu_{2,2})\big) + \cdots + \big(E(\mu_{n-1,1}) + E(\mu_{n-1,2}) + \dots \\ &\qquad\qquad + E(\mu_{n-1,n-2}) + E(\mu_{n-1,n-1})\big) \\ &= r \cdot \big(1 + (2+1) + (3+2+1) + \cdots + (n-1+\cdots+1)\big) \\ &= (n-1) \cdot \left(\binom{2}{2} + \binom{3}{2} + \cdots + \binom{n}{2}\right) = (n-1) \cdot \frac{(n-1)\,n\,(n+1)}{6} \le \frac{n^4}{6} \end{aligned}$$

Substituting this expression into the inequality for the total separation s we get the result. □

Proof of Lemma 6.4.3 Lemma 6.4.3 is proved by a method similar to the one in the proof of Theorem 6.3.2.

Denote $I = \{(i,j), 1 \le i < j \le n\}$. By M denote an uppertriangular matrix $M = (m_{i,j})$, $(i,j) \in I$. We say that i is an *index* and a is a *content* of the matrix

$R(i, i+1; a)$, $a \in \mathbb{F}_q$, $1 \le i < n$. If we apply the generators $R(i_1, i_1+1; a)$, $R(i_2, i_2+1; a)$, ... , the sequence $(i) = (i_1, i_2, \dots)$ is called the *index sequence*.

Let us define an additional structure on τ. In other words, we will "look where we go" and observe not only index of a generator but also whether certain matrix elements are zero or not. To put it more precisely, let us introduce the notion of a *signal* of a walk path.

For a stopping time $\mu_{j,l} = \tau_j + \tau_{j-1} + \cdots + \tau_l$, $1 \le l \le j < n$ we say the signal is green, if the last generator of τ_j was applied to a matrix with a nonzero entry in $(j, j+1)$, if the last generator of τ_{j-1} was applied to a matrix with a nonzero entry in $(j-1, j+1)$, ... , and finally if the last generator of τ_l was applied to a matrix with a nonzero entry in $(l, j+1)$. We say the signal is red otherwise.

For a stopping time $\tau = \mu_{1,1} + \cdots + \mu_{n-1,n-1}$ we say the signal is green if it is green for each of the stopping times $\mu_{j,l}$, $1 \le l \le j < n$. We say it is red otherwise. We can associate with the signal a random variable κ on a space of walk paths $\mathcal{X}$ which takes values 1 if the signal is green, a 0 if the signal is red.

We claim that conditional on the signal being green and the time of the observation, the stopping state of τ is uniformly distributed. In other words, the following identity holds:

$$(*) \qquad P(M = \bar{M} \mid \tau = k, \kappa = 1) = \frac{1}{|G|}$$

This identity immediately implies the lemma. Indeed,

$$\begin{aligned} P(\kappa = 1 \,|\tau = k) = 1 \cdot \left(\left(1 - \frac{1}{q}\right) \cdot 1\right) \cdot \left(\left(1 - \frac{1}{q}\right)^2 \cdot \left(1 - \frac{1}{q}\right) \cdot 1\right) \cdot \dots \\ \cdot \left(\left(1 - \frac{1}{q}\right)^{n-2} \cdot \dots \cdot \left(1 - \frac{1}{q}\right)^2 \cdot \left(1 - \frac{1}{q}\right) \cdot 1\right) \\ = \left(1 - \frac{1}{q}\right)^{\binom{2}{2} + \binom{3}{2} + \cdots + \binom{n-1}{2}} \quad = \quad \left(1 - \frac{1}{q}\right)^{\binom{n}{3}} \end{aligned}$$

Therefore

$$P(\varrho = g | \tau = k) \ge P(\varrho = g | \tau = k, \kappa = 1) \cdot P(\kappa = 1 | \tau = k) = \frac{1}{|G|} \cdot \left(1 - \frac{1}{q}\right)^{\binom{n}{3}}$$

which is exactly what we needed to prove.

Now let us prove the identity $(*)$. Define a linear order "$\prec$" on I as follows:

$$(i, j) \prec (i', j') \text{ if } j < j' \text{ or } j < j',\ i < i'$$

In other words, $(1,2) \prec (1,3) \prec (2,3) \prec (1,4) \prec \cdots \prec (n-2, n) \prec (n-1, n)$.

Now suppose all the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$ are independent and uniform in $\mathbb{F}_q$. We claim that

- after $\mu_{j,l}$, conditioned on the signal being green, all the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$ or $(x, y) = (j, l)$ are independent and uniform in $\mathbb{F}_q$.

The claim follows from the following simple observation. After τ_j we add a nonzero element to the entry in $(j, j+1)$ which is independent of the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$. After τ_{j-1} we add a nonzero element to the entry in $(j-1, j+1)$ which is independent of the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$, etc. After τ_{l+1} we add a nonzero element to the entry in $(l+1, j+1)$ which is independent of the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$. Finally, since the content a of the last generator $R(l, l+1; a)$ is uniform in $\mathbb{F}_q$, after τ_l we add a random element to the entry in $(l, j+1)$ which is independent of the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$.

Now observe that the matrix elements $m_{x,y}$, $(x, y) \prec (j, l)$ remain independent and uniform in $\mathbb{F}_q$. This proves the claim.

From the claim above we have

$$\begin{aligned} P(M = \bar{M} \,|\, (i) = (\bar{i}), \tau = k, \kappa = 1) &= P(m_{1,2} = \bar{m}_{1,2} \,|\, (i) = (\bar{i}), \tau = k, \kappa = 1) \\ \cdot P(m_{1,3} = \bar{m}_{1,3} \,|\, m_{1,2} = \bar{m}_{1,2}, (i) = (\bar{i}), \tau = k, \kappa = 1) &\cdot \ldots \cdot P(m_{n-1,n} = \bar{m}_{n-1,n} | \\ | \, m_{1,2} = \bar{m}_{1,2}, m_{1,3} = \bar{m}_{1,3}, \ldots, m_{n-2,n} = \bar{m}_{n-2,n}, (i) = (\bar{i}), \tau = k, \kappa = 1) & \\ = \left(\frac{1}{q}\right)^{\binom{n}{2}} = \frac{1}{|G|} & \end{aligned}$$

This proves the identity $(*)$ which finishes proof of the lemma. $\square$

Example 6.4.4 Let $n = 3$. By definition, $\tau = \tau_1 + \tau_2 + \tau_1 + \tau_2$. Suppose the index sequence is $(i) = (1, 2, 1, 2)$. Then $\tau = 4$. The condition $\kappa = 1$, means that the content of the second generator is nonzero. Assume it is 1 for convenience. Observe that a matrix

$$M = R(2, 3; c) \cdot R(1, 2; b) \cdot R(2, 3; 1) \cdot R(1, 2; a)$$

is uniformly distributed in $G = U(3; \mathbb{F}_q)$. Indeed,

$$M = \begin{pmatrix} 1 & a+b & b \\ 0 & 1 & 1+c \\ 0 & 0 & 1 \end{pmatrix}$$

which supports the identity $(*)$ in the proof of Lemma 6.4.3.

Now compare these computations with our computations in Example 2.6.5. One can see that Lemma 6.4.3 is an analog of Lemma 2.6.6 for the stopping time τ rather than the stopping time "stop after four steps".

Remark 6.4.5 Note that there are several ways one can refine the construction above. Suppose $\mathbf{A}$ is a stopping call for the stopping time τ (see $S3.1$). The idea of Theorem 3.4.7 is to construct a strong uniform time out of τ which in the notation of the proof of Lemma 6.4.3 will be the following:

- Walk till $\mathbf{A} = 1$. If the signal is green, stop. Else, start all over.

Now, rather than starting all over one can start from the begining of the smaller stopping time $\mu_{j,l}$. This will reduce the power of $\left(1 + \frac{1}{q-1}\right)$ from $O(n^3)$ to $O(n)$.

Another way to improve the bounds is to consider a stopping time τ' which is a maximum of the stopping times $\mu_{j,l}$ rather than sum:

$$\tau' = \max(\mu_{1,1}, \mu_{2,1}, \mu_{2,2}, \dots, \mu_{n-1,n-1})$$

As q tends to infinity this stopping time tends to the strong uniform in a similar sense and we get an asymptotic bound $s = O\big(n^2 \log(n) \log(\log(n))\big)$ which up to a $\log(\log(n))$ factor is tight. These result is due to the author and will be presented elsewhere.

6.5 An affine walk on $\mathbb{F}_q^n$.

Fix a matrix $A \in Mat(n; \mathbb{F}_q)$ and a vector $v \in \mathbb{F}_q^n$ such that vectors v, $A\,v$, ... , $A^{n-1}\,v$ generate the whole vector space $\mathbb{F}_q^n$. Let θ be some real number, $0 \le \theta < 1$. Consider the following Markov chain $\mathcal{M}$:

- Flip a coin with probability of heads θ.
- Choose a random element $a \in \mathbb{F}_q$.
- If tails, $x_{t+1} = x_t + a\,(A^t\,v)$.
- If heads, $x_{t+1} = x_t$.

Consider the following stopping time τ. Start with the zero vector $x_0 = (0)$ and with the zero vector space $W_0 \subset \mathbb{F}_q^n$. At each step we define the vector space $W_{t+1} = W_t$ if heads and $W_{t+1} = \langle W_t, A^t\,v\rangle$ if tails. We stop when $dim(W) = n$, i.e. when $W_t = \mathbb{F}_q^n$. The idea is that at each time our walk is at a random vector in W even conditioned that we know W and the time t.

Algorithm 6.5.1 Set $x_0 = (0)$, $W_0 \subset \mathbb{F}_q^n$.

- Flip a coin with probability of heads θ.
- Choose a random element $a \in \mathbb{F}_q$.
- If tails, $x_{t+1} = x_t + a \cdot (A^t\,v)$; $W_{t+1} = \langle W_t, A^t\,v\rangle$;
- If heads, $x_{t+1} = x_t$.
- If $dim(W) = n$, stop. Else, return to the beginning.

Theorem 6.5.2 The stopping time τ defined by the Algorithm 6.5.1 is strong uniform.

Proof We claim that at step t we have

$$P(x_t = \bar{v} | W_t = \bar{W}) = \begin{cases} \dfrac{1}{q^{dim(\bar{W})}}, & \bar{v} \in \bar{W} \\ 0, \quad \text{otherwise} \end{cases}$$

The claim is obvious by induction. Indeed, if w is a uniformly distributed vector in W_t, then $w + a \cdot A^t\,v$ is uniformly ditsributed in $\langle W_t, A^t\,v\rangle$ since a is uniformly distributed in $\mathbb{F}_q$.

Since we stop when $W = \mathbb{F}_q^n$, the claim implies

$$P(\varrho = \bar{v} | \tau = k) = P(\varrho = \bar{v} | W_\tau = \mathbb{F}_q^n) = \frac{1}{q^n}$$

This finishes the proof. □

Theorem 6.5.3 Let $A^n = id$. Then the stopping time τ is perfect and the separation distance satisfies

$$s_{kn+i} = 1 - (1-\theta^k)^{n-i}\,(1-\theta^{k+1})^i$$

for any $i, k \in \mathbb{Z}_+$, $0 \le i \le n-1$.

Proof Consider the vector $\widehat{v} = v + A\,v + A^2\,v + \cdots + A^{n-1}\,v$. We claim that $\widehat{v}$ is an extremal element, i.e. whenever the chain gets to $\widehat{v}$ it stops there. Indeed, since $A^{kn+i}\,v = A^i\,v$ for any $0 \le i \le n-1$, in order to get to $\widehat{v}$ one has to get tails at least once for every number i modulo n. This means that then $W = \mathbb{F}_q^n$ and the chain must stop in $\widehat{v}$. This proves the first part of the Theorem.

By Theorem 6.5.2 we have $s_t = P(\tau > t) = 1 - P(\tau \le t)$. After $t = kn+i$ steps, $0 \le i \le n-1$, the probability of W_t containing $A^j\,v$, $0 \le j \le n-1$, is equal to $(1-\theta^k)$ if $j \ge i$ and $(1-\theta^{k+1})$ if $j < i$. Since these events are independent, we conclude

$$s_{kn+i} = 1 - (1-\theta^k)^{n-i}\,(1-\theta^{k+1})^i$$

which finishes the proof of the second part. □

Example 6.5.4 Let $q = 2$, $v = (0, \ldots, 0, 1)$, $\theta = 0$, and A is the following matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Observe that $\langle v, A\,v, \ldots, A^{n-1}\,v\rangle = \mathbb{F}_2^n$. We can think of the vectors we get as of the coin flipping outcomes. Of course, here $s = n$. This problem is similar to the problem considered in [A1].

Example 6.5.5 Let $q = 2$, $v = (x_1, \ldots, x_{n-1}, 1)$, and let B be the following nonsingular matrix

$$B = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 1 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

In a paper [DG] Diaconis and Graham consider the following a Markov chain $\mathcal{M}'$:

- Flip a coin with probability of heads θ'.
- If tails, $y_{t+1} = B\,y_t + v$.
- If heads, $y_{t+1} = B\,y_t$.

Let $\theta \ge \frac{1}{2}$, $\theta = 2\,\theta' - 1$, $A = B^{-1}$. Let $\mathcal{M}$ be the Markov chain defined in the beginning of the section. We first show that the sepration distances for the Markov chain $\mathcal{M}'$ is equal to the sepration distances for $\mathcal{M}$. Indeed, define $x_t = A^{t-1}\,y_t$. We have

$$y_{t+1} = B\,y_t + \epsilon \cdot v$$

where $\epsilon = 0$ with probability θ' and $\epsilon = 1$ with probability $1 - \theta'$. Rewrite the last equality in terms of x_t:

$$B^t\, x_{t+1} = B \cdot B^{t-1}\, x_t + \epsilon \cdot v$$

Multiplying by A^t on the left we get:

$$x_{t+1} = x_t + \epsilon \cdot A^t\, v$$

Since $\theta' = \frac{1+\theta}{2} = \theta + \frac{1-\theta}{2}$, this means that the Markov chain x_t is isomorphic to $\mathcal{M}$.

Denote by Q^k, R^k the probability distributions of the chains $\mathcal{M}$, $\mathcal{M}'$. From above for every vector $z \in \mathbb{F}_q^n$

$$Q^k(z) = R^k(A^{t-1}\, z)$$

By definition, for any $k > 0$ the separation distance s'_k of the chain $\mathcal{M}'$ is equal to the total separation s_k of the chain $\mathcal{M}$. Here is how we can compute the latter.

Observe that when $n = 2\,m$ we have $A^n = id$. Therefore in this case, by Theorem 6.5.3, we also have

$$s_{kn+i} = 1 - (1 - \theta^k)^{n-i}\,(1 - \theta^{k+1})^i$$

for any $i, k \in \mathbb{Z}_+$, $0 \le i \le n-1$. In particular, when $\theta' = \frac{1}{2}$, $\theta = 0$, and $s_1 = s_2 = \cdots = s_{n-1} = 1$, $s_n = 0$.

References

[A1] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, Lecture Notes in Mathematics **986** (1983).

[A2] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, J. Theor. Probability **2** (1980), 91–100.

[AD1] D. Aldous, P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly **93** (1986), 333–348.

[AD2] D. Aldous, P. Diaconis, *Strong uniform times and finite random walks*, Advances in Applied Math. **8** (1987), 69–97.

[AF] D. Aldous, J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, monograph in preparation, 1996.

[ASE] N. Alon, J. H. Spencer, P. Erdős, *The Probabilistic Method*, Wiley, New York, 1992.

[Ar] A. D'Aristotile, *The nearest neighbor random walk on subspaces of a vector space and rate of convergence*, J. Theoretical Probability **8** (1995), 321–346.

[AP] A. Astashkevich, I. Pak, *Random walks on nilpotent and supersolvable groups*, in preparation (1996).

[Babai] L. Babai, *Automorphism groups, isomorphism, reconstruction*, in Handbook of Combinatorics (R. L. Graham, M. Groetschel, and L. Lovasz, eds.) (1996), Elsevier.

[BS] L. Babai, Á. Seress, *On the diameter of Cayley graphs of the symmetric group*, J. Comb. Theory (A) **49** (1988), 175–179.

[BE] H. Bateman, A. Erdélyi, *Higher Transcedential Functions, Volume 1*, Mc Graw-Hill, New York, NY, 1953.

[Bel] E. Belsley, *Random walks on distance regular graphs* (1996), in preparation.

[Bou] N. Bourbaki, *Groupes et algèbres de Lie, Ch I, IV, V, VI, VII, VIII*, Hermann, Paris, 1960, 1972.

[CW] E. Calabi, H. Wilf, *On the sequential and random selection of subspaces over a finite field*, J. Comb. Theory (A) **22** (1977), 107–109.

[Che] I.V. Cherednik, *A new interpretation of Gelfand-Tzetlin bases*, Duke Math. J. **54** (1987), 563–571.

[Chung] F. R. K. Chung, *Spectral Graph Theory* (Regional conference series in mathematics, no 92), American Mathematical Society, Providence, RI, 1997.

[CM] H.S.M. Coxeter, W.O.J. Moser, *Generators and relations for discrete groups* (third edition), Springer, Berlin, 1972.

[CDS] D.M. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs,* Second edition, Academic Press, New York, 1980.

[Dem] M. Demazure, *Désingularisation des variétes de Schubert généralisées*, Ann. Sci. Éc. Norm. Sup. **7** (1974), 53–88.

[D] P. Diaconis, *Group Representations in Probability and Statistics*, IMS, Hayward, California, 1988.

[Dm] P. Diaconis, *A group theoretic interpretation of the records to cycles map* (1995), unpublished manuscript.

[Dc] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. U.S.A. **93** (1996), 1659–1664.

[DP] P. Diaconis, personal communication.

[DF] P. Diaconis, J. A. Fill, *Strong stationary times via new form of duality*, The Annals of Probability **18** (1990), 1483–1522.

[DG] P. Diaconis, R. Graham, *An affine walk on the hypercube*, J. Comp. Appl. Math **41** (1992), 215–235.

[DGM] P. Diaconis, R. Graham, J. Morrison, *Asymptotic analysis of a random walk on a hypercube with many dimensions*, Random structures and algorithms **1** (1990), 51–72.

[DSC] P. Diaconis, L. Saloff–Coste, *Comparison techniques for random walk on finite groups*, The Annals of Probability **21** (1993), 2131–2156.

[D-S-C] P. Diaconis, L. Saloff–Coste, *Moderate grows and random walk on finite groups*, Geom. Funct. Anal. **4** (1994), 1–36.

[DSh1] P. Diaconis, M. Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahr. verw. Gebiete **57** (1981), 159–179.

[DSh2] P. Diaconis, M. Shahshahani, *Time to reach stationarity in the Bernoulli-Laplace diffusion model*, SIAM J. Math'l Analysis **18** (1987), 206–218.

[DSh3] P. Diaconis, M. Shahshahani, *The subgroup algorithm for generating uniform random variables*, Prob. in Eng. and Info. Sci. **1** (1987), 15–32.

[DM] J. Dixon, B. Mortimer, *Permutation Groups*, Springer, New York, 1996.

[Ell] J. Ellenberg, *A sharp diameter bound for an upper triangular matrix group*, Senior Honor Thesis, Harvard U. (1995).

[EKK] V.A. Emelichev, M.M. Kovalev, M.K. Kravtsov, *Polytopes, Graphs and Optimization*, Cambridge University Press, New York, 1984.

[Feller] W. Feller, *An introduction to Probability theory and its applications* (third edition), John Wiley, New York, 1970.

[Fill] J. A. Fill, *An exact formula for move-to-front rule for self-organizing lists*, J. Theor. Prob. **9** (1996), 113–160.

[FOW] L. Flatto, A. M. Odlyzko, D. B. Wales, *Random shuffles and group representations*, Ann. Prob. **13** (1985), 155–178.

[Ful] J. Fulman, *Transvections* (1995), preprint.

[GR] I.M. Gelfand, V.S. Retakh, *A theory of noncommutative determinants and characteristic functions of graphs*, Funct. Anal. Appl. **26** (1992), 1–20.

[Gluck] D. Gluck, *Characters and random walks on finite classical groups*, Advances in Mathematics, to appear (1995).

[Gor] D. Gorenstein, *Finite Simple Groups*, Plenum, New York, 1982.

[GJ] I. P. Goulden, D. M. Jackson, *Combinatorial enumeration*, John Wiley, New York, 1983.

[Gr] A. Greenhalgh, *Random walks on groups with subgroup invariance properties, Ph. D. dissertation*, Stanford U., 1987.

[Hall] M. Hall, *The Theory of Groups*, Chelsea, New York, NY, 1976.

[HJ] S. Handjani, D. Jungreis, *Rate of convergence for shuffling cards by transpositions*, J. Theoret. Probability **4** (1996), 983–993.

[Hild] M. Hilderbrand, *Generating random elements in $SL_n(\mathbb{F}_q)$ by random transvections*, J. Alg. Combinatorics **1** (1992), 133–150.

[H1] J. Humphreys, *Linear algebraic groups*, Springer, Berlin, 1975.

[H2] J. Humphreys, *Reflection groups and Coxeter groups*, Cambridge University Press, Cambridge, UK, 1990.

[Jucys] A. Jucys, *On the Young operators of the symmetric group* (in Russian), Lietuvos Fizikos Rinkinys **6** (1966), 163–180.

[KPP] A. G. Kuznetsov, I. M. Pak, A. E. Postnikov, *Increasing trees and alternating permutations*, Russian Math. Survey **49** (1994), 79–114.

[Lall] G. Lallement, *Semigroups and Combinatorial Applications*, Wiley, New York, NY, 1979.

[Lang] S. Lang, *Algebra*, Addison-Wesley, Reading, MA, 1965.

[Lulov] N. Lulov, *Random Walks on the Symmetric Group Generated by Conjugacy Classes* (PH. D. thesis), Harvard University, 1996.

[LP] N. Lulov, I. Pak, *Random walks on the symmetric group generated by cycles*, preprint (1996).

[Mac] I. G. Macdonald, *Notes on Schubert Polynomials*, Department de mathematiques et d'informatique Université du Québec, Montréal, 1991.

[MO] A. Marshall, I. Olkin, *Multivariate distributions generated from mixtures of convolution and product families* IMS Lecture Notes Monograph Ser., 16, Inst. Math. Statist., Hayward, CA, 1990.

[Matt1] P. Matthews, *A strong uniform time for random transpositions*, J. Theoretical Probability **1** (1988), 411–423.

[Matt2] P. Matthews, *Some sample path properties of a random walk on the cube*, J. Theoretical Probability **2** (1989), 129–146.

[Matt3] P. Matthews, *Strong stationary times and eigenvalues*, J. Applied Probability **29** (1992), 228–233.

[N] M. Nazarov, *Young's Symmetrizers for Projective Representations of the Symmetric Group*, RIMS Preprint No 900, Kyoto University, Kyoto, Japan, 1992.

[Odl] A. Odlyzko, *Asymptotic Enumeration Methods*, in Handbook of Combinatorics, vol. 2 (R. L. Graham, M. Groetschel, and L. Lovasz, eds.) (1996), Elsevier.

[P] I. Pak, *When and how n choose k,* preprint (1996).

[P1] I. Pak, *Problème des ménages, rook placements and random walk on the symmetric group,* manuscript in preparation (1997).

[P2] I. Pak, *Reduced decompositions of permutations in terms of star transpositions an enumeration of trees,* manuscript in preparation (1997).

[PPR] I. Pak, A. Postnikov, V. Retakh, *Noncommutative Lagrange Inversion,* preprint (1995).

[Ran] D. Randall, *Efficient random generation of nonsingular matrices*, Random Structures and Algorithms **4** (1993), 111–118.

[Rior] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.

[Roi] Y. Roichman, *Upper bound on characters of the symmetric groups*, Invenciones to appear (1995).

[Serre] J. P. Serre, *Représentations Linéaires des Groupes Finis*, Hermann, Paris, 1967.

[Sh] A. N. Shiryaev, *Probability* (in Russian), Nauka, Moscow, 1980.

[Stan] R. P. Stanley, *Enumerative Combinatorics,* Vol. 1, Wadsworth & Brooks/Cole, California, 1986.

[Stong1] R. Stong, *Random walk on the upper triangular matrices*, Ann. Prob. **23** (1995), 1939–1949.

[Stong2] R. Stong, *Eigenvalues of random walks on groups*, Ann. Prob. **23** (1995), 1961–1981.

[Vishne] U. Vishne, *Mixing and covering in the symmetric group*, preprint (1996).

[VDW] B. L. van der Waerden, *Algebra I, II*, Springer, Berlin, 1971.

[Wein] M. Weinstein, *Examples of Groups*, Polygonal Publishing, Passaic, NJ, 1977.

[W] H. Weyl, *Classical Groups*, Princeton University Press, Princeton, NJ, 1939.

[WW] E. T. Whittaker, G. N. Watson, *A Course of Modern Analysis* (Fourth Edition), Cambridge University Press, Cambridge, UK, 1927.

[Z] G. Ziegler, *Lectures on Polytopes, Graduate Texts in Mathematics 152*, Springer, New York, 1995.