Bachelorarbeit

Measure Concentration for Symplectic Groups

Florian Starke

October 5, 2018

Technische Universität Dresden
Fakultät Mathematik
Institut für Geometrie

Betreuender Hochschullehrer: Prof. Dr. Andreas Thom

Contents

Preface

If we have a finite group of matrices, then we can equip it with the rank metric and the normalized Haar measure to obtain a *metric measure space*. There is a well defined limit of ... Carderi and Thom showed in [**?**] that the limit of SL_n is *extremely amenable*. The goal of this thesis is to generalize this result to limits of other matrix group families, namely unitary, symplectic, and orthogonal matrices. The general strategy will be the following: given a family $(G_n)_{n\in\mathbb{N}}$ of (mm) matrix groups we first find an upper bound for the *concentration function* of G_n using a consequence of Azemas inequality [**?**]. As the upper bounds converge to zero we conclude that $(G_n)_{n\in\mathbb{N}}$ is a *Lévy family*, making their limit a *Lévy group*. Finally, we know from [**?**] that every Lévy group is extremely amenable.

1 Introduction

Define limit of G_n

Examples of matrices in the limit

structure of thesis:

1. Azema
2. Thoms proof (matrices as automorphisms but without form)
3. want to generalize this so we need a form Hence extending the automorphism becomes harder so use Witts lemma
4. generalized version of the proof
5. application coloring theorem

2 Limits of matrix groups and extreme amenability

Let $\mathrm{GL}_n(q)$ be the general linear group over the q element field $\mathbb{F}_q$ and let G be a subgroup of $\mathrm{GL}_n(q)$. We can equip G with the (normalized) *rank-metric* $d(g,h) := \frac{1}{n}\,\mathrm{r}(g-h)$. Since all matrices in G have full rank, this metric is bi-invariant, i.e. $d(kg,kh) = d(g,h) = d(gk,hk)$ for all $g,h,k\in G$. Let $G_n \leq \mathrm{GL}_{2^n}(q)$ be a family of subgroups, such that $\begin{pmatrix} g & 0 \\ 0 & g \end{pmatrix} \in G_{n+1}$ for all $g \in G_n$. Note that the map

$$\varphi_n\colon G_n \mapsto G_{n+1}, \text{ where } \varphi_n(g) = \begin{pmatrix} g & 0 \\ 0 & g \end{pmatrix}$$

is an isometric homomorphism for all $n\in\mathbb{N}$. Hence we can define the inductive limit of $(G_n)_{n\in\mathbb{N}}$. We denote the metric completion of this limit by $\mathrm{clim}_{n\to\infty} G_n$.

Lemma 1. *The group* $\mathrm{clim}_{n\to\infty} G_n$ *is a topological group.*

Proof. The bi-invariance of d is preserved by the limit and the completion. ... $\square$

Now that we have a topology on $G := \operatorname{clim}_{n\to\infty} G_n$ we can ask whether it is *extremely amenable*, i.e. every continuous action of G on a compact topological space admits a fixed point. It is hard to show this directly, but we know that every Lévy group is extremely amenable. Hence we will show that for suitable $(G_n)_{n\in\mathbb{N}}$ the limit G will be a Lévy group.

Before we can define Lévy groups we need the following definition.

Definition 2. A *metric measure space* (mm-space) $\boldsymbol{X}$ is a triple (X, d, μ), where d is a metric on the set X and μ is a measure on the Borel σ-algebra induced by d. We will always assume that $\mu(X) = 1$. For any set $A \subseteq X$ denote the $r-$*neighborhood* of A, i.e. $\{x \in X \mid \exists y \in A.\, d(x,y) < r\}$, by $N_r(A)$. The *measure concentration function* of $\boldsymbol{X}$ is defined as

Florian sagt: "definition ugly, use d_A?"

$$\alpha_{\boldsymbol{X}}(r) = \sup\{1 - \mu(N_r(A)) \mid A \subseteq X, \mu(A) \geq \frac{1}{2}\}.$$

A family of mm-spaces $\boldsymbol{X}_n$ with diameter 1 is called a *Lévy family* if

$$\alpha_{\boldsymbol{X}_n}(r) \to 0$$

for all $r > 0$.

A topological space $\boldsymbol{X}$ is a *Polish space* if it is homeomorphic to a complete metric space that has a countable dense subset.

Now we can come back to groups.

Definition 3. A *Polish group* G is a topological group where the underlying topological space is a Polish space. A *Lévy group* is a group G equipped with a metric d, where

- G with the topology induced by d is a Polish group and
- there is a sequence $(G_n)_{n\in\mathbb{N}}$ of compact subgroups, such that $(G_n, d|_{G_n}, \mu_n)_{n\in\mathbb{N}}$ is a Lévy family. Here μ_n is the normalized Haar measure of G_n.

Florian sagt: "lim G_n dense in G?"

Note that the normalized Haar measure of G_n is just the normalized counting measure. The following theorem from [**?**] gives the desired connection to extreme amenability.

Theorem 4. *Every Lévy group is extremely amenable.*

To apply this theorem to our setting we need the following lemma.

Lemma 5. *Let $G_n \leq \mathrm{GL}_{2^n}(q)$ and $G = \operatorname{clim}_{n\to\infty} G_n$. Then G is a Polish group.*

Proof. By Lemma 1 G is already a topological group and by definition it is also a complete metric space. Furthermore, every G_n is finite. Hence the inductive limit of the G_n is a countable dense subset of G. $\square$

Whether G is also a Lévy group depends on the particular choice of $(G_n)_{n\in\mathbb{N}}$. To show that for certain sequences G will be a Lévy group, we will bound $\alpha_{G_n}(r)$. The next section develops the methods necessary to obtain this upper bound.

3 An upper bound for the measure concentration function

In this section we will prove Azema's inequality and as a consequence, we will obtain an upper bound for the measure concentration function. As the next results rely heavily on stochastic methods we will briefly introduce the necessary notions. Since the G_n are all finite and equipped with the normalized counting measure we will only consider *probability spaces* (X, Σ, μ), where X is finite, Σ is a σ-algebra over X, and $\mu(A) = |A|/|X|$ for $A \subseteq X$. Most of the statements in this section hold in a more general setting **[?]**. Note that Σ has a very nice representation.

Lemma 6. *Let Σ be a σ-algebra over a finite set X, then Σ is the smallest σ-algebra containing the partition $A_1, \ldots, A_n$, where the A_i's are the minimal nonempty sets in Σ.*

Proof. First we show that $A_1, \ldots, A_n$ is a partition of X. Since $A_i \cap A_j \in \Sigma$ we conclude, by minimality of A_i and A_j, that either $i = j$ or $A_i \cap A_j = \emptyset$. Clearly, every element of X is contained in a one of the A_i.

For $A \in \Sigma$ we have, again by minimality, that $A \cap A_i$ is either A_i or $\emptyset$. Therefore A can be written as a union of A_i's. $\square$

Note that it follows from the proof that any $A \in \Sigma$ can be written as $\bigcup_{i \in I} A_i$ for a suitable I. This lemma allows us to use partitions and σ-algebras interchangeably. We will denote the partition corresponding to Σ by $A_1, \ldots, A_n$, for Σ' we will use $A'_1, \ldots, A'_{n'}$, etc. The next definition is simplified a lot by only considering finite X.

Definition 7. Let (X, Σ, μ) be a probability space, $f\colon X \to \mathbb{R}$ be a measurable function, and Σ' be a sub-σ-algebra of Σ. Then the *conditional expectation* of f with respect to Σ' is defined as

$$\mathbb{E}(f \mid \Sigma') := \sum_{i=1}^{n'} \mathbb{E}(f \mid A'_i) \cdot \mathbb{1}_{A'_i}.$$

One often thinks of Σ' as the available information, a finer partition means more information. The conditional expectation $\mathbb{E}(f \mid \Sigma')$ is the best approximation of f given only the information from Σ'. With this intuition the statements from the following lemma are not surprising.

Lemma 8. *Let (X, Σ, μ) be a probability space, $f, g\colon X \to \mathbb{R}$ be measurable functions, $\Sigma'' \subseteq \Sigma' \subseteq \Sigma$ be sub-σ-algebras. Then*

i) if $f \le g$, then $\mathbb{E}(f \mid \Sigma') \le \mathbb{E}(g \mid \Sigma')$,

ii) for any Σ'-measurable function $h\colon X \to \mathbb{R}$ we have $\mathbb{E}(hf \mid \Sigma') = h \cdot \mathbb{E}(f \mid \Sigma')$,

iii) also $\mathbb{E}(\mathbb{E}(f \mid \Sigma') \mid \Sigma'') = \mathbb{E}(f \mid \Sigma'') = \mathbb{E}(\mathbb{E}(f \mid \Sigma'') \mid \Sigma')$.

Proof. To i): If $f \leq g$, then

$$\mathbb{E}(f \mid \Sigma') = \sum_{i=1}^{n} \mathbb{E}(f \mid A_i') \cdot \mathbb{1}_{A_i'} \leq \sum_{i=1}^{n} \mathbb{E}(g \mid A_i') \cdot \mathbb{1}_{A_i'} = \mathbb{E}(g \mid \Sigma').$$

To ii): Let $h \colon X \to \mathbb{R}$ be Σ'-measurable function, then $h = \sum_{i=1}^{n'} h_i \mathbb{1}_{A_i'}$. Now

$$\begin{aligned}
\mathbb{E}(hf \mid \Sigma') &= \sum_{i=1}^{n'} \mathbb{E}(hf \mid A_i')\mathbb{1}_{A_i'} \\
&= \sum_{i=1}^{n'} h_i \mathbb{E}(f \mid A_i')\mathbb{1}_{A_i'} \\
&= h \cdot \mathbb{E}(f \mid \Sigma').
\end{aligned}$$

To iii): Note that $\mathbb{E}(\mathbb{E}(f \mid \Sigma') \mid A') = \mathbb{E}(f \mid A')$ for all $A' \in \Sigma'$.

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(f \mid \Sigma') \mid \Sigma'') &= \sum_{i=1}^{n''} \mathbb{E}(\mathbb{E}(f \mid \Sigma') \mid A_i'') \cdot \mathbb{1}_{A_i''} \\
&= \sum_{i=1}^{n''} \mathbb{E}(f \mid A_i'') \cdot \mathbb{1}_{A_i''} && (A_i'' \in \Sigma') \\
&= \mathbb{E}(f \mid \Sigma'') \\
&= \sum_{j=1}^{n''} \mathbb{E}(f \mid A_j'') \cdot \mathbb{1}_{A_j''} \cdot \sum_{i=1}^{n'} \mathbb{1}_{A_i'} \\
&= \sum_{i=1}^{n'} \sum_{j=1}^{n''} \mathbb{E}(\mathbb{E}(f \mid A_j'') \mid A_i') \cdot \mathbb{1}_{A_j''} \cdot \mathbb{1}_{A_i'} \\
&= \sum_{i=1}^{n'} \sum_{j=1}^{n''} \mathbb{E}(\mathbb{E}(f \mid A_j'') \cdot \mathbb{1}_{A_j''} \mid A_i') \cdot \mathbb{1}_{A_i'} && \text{(by ii)} \\
&= \mathbb{E}(\mathbb{E}(f \mid \Sigma'') \mid \Sigma')
\end{aligned}$$

This concludes the proof. □

The following lemma might not seem very interesting, but changing the exponent from x to x^2 is the very foundation for Azema's inequality.

Lemma 9. *For all* $x \in \mathbb{R}$

$$e^x \leq x + e^{x^2}.$$

Proof. Note that for $x = 0$ both sides are equal to 1. As both sides are differentiable it suffices to show that the derivative of the right hand side is larger than the derivative of the left hand side for all $x \geq 0$ and smaller for all $x \leq 0$. Hence, we want to show

$$e^x \geq 1 + 2xe^{x^2} \text{ for all } x \leq 0 \qquad \text{and} \qquad e^x \leq 1 + 2xe^{x^2} \text{ for all } x \geq 0.$$

As for $x = 0$ both sides are again equal to 1 we can reduce the problem, by similar reasoning, to the question whether

$$e^x \leq 2e^{x^2} + 4x^2 e^{x^2} \text{ for all } x \in \mathbb{R}.$$

- For $x = 0$ the terms reduce to $1 \leq 2$.
- For $x < 0$ the left hand side is bounded by 1, while the right hand side is still larger that 2.
- For $1 \leq x$ we have $x \leq x^2$ and the inequality holds trivially.
- For $0 < x < 1$ note that the both sides are increasing. Hence the inequality holds for all x with $e^x \leq 2$. Finally, $\ln 2 \geq \frac{1}{2}$ and therefore the right hand side with $x = \ln 2$ evaluates to a number larger then e.

$\square$

Before we will prove Azema's inequality let us introduce some useful notation. Whenever there is no danger of confusion we will abbreviate sets of the form

$$\{x \in X \mid \text{Condition}(x) \text{ holds}\} \qquad \text{by} \qquad \{\text{Condition}\}.$$

For example $\{x \in X \mid f(x) = c\}$ becomes $\{f = c\}$.

Lemma 10. *[Azema's inequality] Let (X, Σ, μ) be a probability space, $f \colon X \to \mathbb{R}$ a measurable function, and $\{X\} = \Sigma_0 \subseteq \cdots \subseteq \Sigma_n = \Sigma$ a chain of sub-σ-algebras. Define $f_i := \mathbb{E}(f \mid \Sigma_i)$ and $d_i := f_i - f_{i-1}$. Then for every $\varepsilon \geq 0$*

$$\mu(\{|f - \mathbb{E}(f)| \geq \varepsilon\}) \leq 2 \cdot \exp\left(-\frac{\varepsilon^2}{4 \cdot \sum_{i=1}^{n} \|d_i\|_\infty^2}\right).$$

Note that $(f_i, \Sigma_i)_{0 \leq i \leq n}$ is a discrete martingale.

Proof. First, observe that $f_0 = \mathbb{E}(f \mid \{X\}) = \mathbb{E}(f)$ and $f_n = \mathbb{E}(f \mid \Sigma) = f$. Using a simple telescoping sum we obtain $f - \mathbb{E}(f) = d_1 + \cdots + d_n$. Therefore

$$\begin{aligned}
\mu(\{f - \mathbb{E}(f) \geq \varepsilon\}) &= \mu(\{\sum_{i=1}^{n} d_i \geq \varepsilon\}) \\
&= \mu(\{\lambda \cdot \sum_{i=1}^{n} d_i \geq \lambda\varepsilon\}) && \text{(for } \lambda > 0\text{)} \\
&= \mu(\{e^{\lambda \cdot \sum_{i=1}^{n} d_i - \lambda\varepsilon} \geq 1\}) \\
&\leq \mathbb{E}(e^{\lambda \cdot \sum_{i=1}^{n} d_i}) \cdot e^{-\lambda\varepsilon} && (*) \\
&= \mathbb{E}(e^{\lambda d_1} \cdot \ldots \cdot e^{\lambda d_{n-1}} \cdot \mathbb{E}(e^{\lambda d_n} \mid \Sigma_{n-1})) \cdot e^{-\lambda\varepsilon} && \text{(Lemma 8)} \\
&\leq \mathbb{E}(e^{\lambda d_1} \cdot \ldots \cdot e^{\lambda d_{n-1}}) \cdot e^{\lambda^2 \cdot \|d_n\|_\infty^2} \cdot e^{-\lambda\varepsilon} && (**) \\
&\;\;\vdots \\
&\leq e^{\lambda^2 \cdot \|d_1\|_\infty^2} \cdot \ldots \cdot e^{\lambda^2 \cdot \|d_{n-1}\|_\infty^2} \cdot e^{\lambda^2 \cdot \|d_n\|_\infty^2} \cdot e^{-\lambda\varepsilon} \\
&= e^{\lambda^2 \cdot \sum_{i=1}^{n} \|d_i\|_\infty^2 - \lambda\varepsilon}.
\end{aligned}$$

For $(*)$ note that for any measurable function $g\colon X \to \mathbb{R}$ with $g \geq 0$ we have

$$\mu(\{g \geq 1\}) = \mathbb{E}(\mathbb{1}_{\{g \geq 1\}}) \leq \mathbb{E}(g).$$

For $(**)$ we need to use Lemma 9

$$\begin{aligned}
\mathbb{E}(e^{\lambda d_i} \mid \Sigma_{i-1}) &\leq \mathbb{E}(\lambda d_i \mid \Sigma_{i-1}) + \mathbb{E}(e^{\lambda^2 d_i^2} \mid \Sigma_{i-1}) \\
&= \lambda \cdot \mathbb{E}(f_i - f_{i-1} \mid \Sigma_{i-1}) + \mathbb{E}(e^{\lambda^2 d_i^2} \mid \Sigma_{i-1}) \\
&\leq 0 + e^{\lambda^2 \|d_i\|_\infty^2}. && \text{(Lemma 8)}
\end{aligned}$$

Substituting $-\frac{\varepsilon^2}{\sum_{i=1}^{n} \|d_i\|_\infty^2}$ for λ we conclude that

$$\mu(\{f - \mathbb{E}(f) \geq \varepsilon\}) \leq \exp\left(-\frac{\varepsilon^2}{4 \cdot \sum_{i=1}^{n} \|d_i\|_\infty^2}\right).$$

The same calculations with $-d_i$ instead of d_i yield the dual inequality

$$\mu(\{f - \mathbb{E}(f) \leq -\varepsilon\}) \leq \exp\left(-\frac{\varepsilon^2}{4 \cdot \sum_{i=1}^{n} \|d_i\|_\infty^2}\right).$$

These two statements obviously give us the desired result. □

Since μ is the counting measure Azema's inequality bounds the number of elements for which f differs more than ε from its mean. This seems at least somewhat connected to the measure concentration function, as there we want to show that for any set A with $\mu(A) \geq \frac{1}{2}$ only a few elements are more than ε away from A. The next goal is to formalize this connection To achieve this we first need to introduce a new property of mm-spaces.

Definition 11. Let $\boldsymbol{X} = (X, d, \mu)$ be a finite mm-space. The *length* of $\boldsymbol{X}$ is the minimum over all l with the following property. There is a refining sequence of partitions

$$\{X\} = \Omega_0 \prec \cdots \prec \Omega_n = \{\{x\} \mid x \in X\},$$

where for every $i \in \{1, \ldots, n\}$ there is an a_i such that $\sum_{i=1}^n a_i^2 = l^2$ and for every $A \in \Omega_{i-1}$, $x, y \in A$ there is an isomorphism (of metric spaces) $\phi\colon [x]_i \to [y]_i$ with

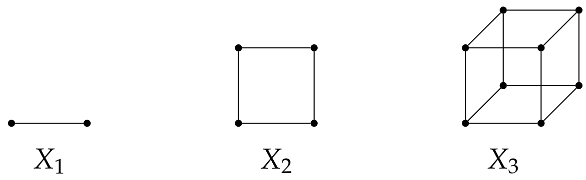
$$d(z, \phi(z)) \leq a_i \text{ for all } z \in [x]_i.$$

Note that since μ is the counting measure ϕ is also an isomorphism of mm-spaces. As this definition is quite hard we will look at some properties and examples of the length of $\boldsymbol{X}$ before proceeding.

Lemma 12. *Let $\boldsymbol{X}$ be a finite mm-space. Then the length of $\boldsymbol{X}$ is at most the diameter of $\boldsymbol{X}$.*

Proof. Consider only the two partitions $\{X\} \prec \{\{x\} \mid x \in X\}$. □

Example 13. Let us look at the n-dimensional cube $X_n = \{0,1\}^n$.



We will only consider the following sequence of partitions

$$\Omega_0 \prec \cdots \prec \Omega_n \text{ with } \Omega_i = \{wX_{n-i} \mid w \in \{0,1\}^i\}.$$

First, we equip X_n with the euclidean metric and rescale it such that the diameter is 1. To bound the length of the resulting space $\boldsymbol{X}_n^E$ consider $[x]_i \neq [y]_i$. Note that x and y are w.l.o.g. of the form $w0u$ and $w1v$ for some $w \in \{0,1\}^{i-1}$, $u, v \in \{0,1\}^{n-i}$. The isomorphism ϕ takes an element $w0u'$ in $[x]_i$ and maps it to $w1u'$. The length of a side in $\boldsymbol{X}_n^E$ is $\frac{1}{\sqrt{n}}$, hence every a_i is $\frac{1}{\sqrt{n}}$ for every i and the length of $\boldsymbol{X}_n^E$ is bounded by $\sum_{i=1}^n \frac{1}{\sqrt{n}^2} = 1$.

Secondly, we use the hemming metric and obtain the mm-space $\boldsymbol{X}_n^H$ with diameter 1. It has side length $\frac{1}{n}$ and therefore the length of $\boldsymbol{X}_n^H$ is bounded by $\sum_{i=1}^n \frac{1}{n^2} = \frac{1}{n}$. We see that here the length of $\boldsymbol{X}_n^H$ converges to 0 as n tends to infinity. We will show that this means that the measure concentration function $\alpha_{\boldsymbol{X}_n^H}(r)$ also goes to 0 for any fixed $r > 0$.

Lemma 14. *Let $\boldsymbol{X} = (X, d, \mu)$ be a finite mm-space of length l and $f\colon X \to \mathbb{R}$ be a 1-Lipschitz function. Then*

$$\mu(\{|f - \mathbb{E}(f)| \geq \varepsilon\}) \leq 2 \exp\left(-\frac{\varepsilon^2}{4l^2}\right) \textit{ for every } \varepsilon > 0.$$

Proof. Let $\Omega_0 \prec \cdots \prec \Omega_n$ be a refining sequence of partitions with $a_1, \ldots, a_n$ as in Definition 11 such that $\sum_{i=1}^n a_i^2 = l^2$. These partitions correspond to σ-algebras $\Sigma_0 \subseteq \cdots \subseteq \Sigma_n$. Now we can apply Azema's inequality to obtain

$$\mu(\{|f - \mathbb{E}(f)| \geq \varepsilon\}) \leq 2 \cdot \exp\left(-\frac{\varepsilon^2}{4 \cdot \sum_{i=1}^n \|d_i\|_\infty^2}\right),$$

where $f_i = \mathbb{E}(f \mid \Sigma_i)$ and $d_i = f_i - f_{i-1}$ as before. Therefore we only need to show that $\|d_i\|_\infty \leq a_i$. Since on any $A \in \Omega_{i-1}$ we have $f_{i-1} = \mathbb{E}(f_i \mid A)$ it suffices to show that for all $A \in \Omega_{i-1}$ it holds that $f_i(x) - f_i(y) \leq a_i$ for all $x, y \in A$. Let $\phi \colon [x]_i \to [y]_i$ be the isomorphism from Definition 11.

$$\begin{aligned} f_i(x) - f_i(y) &= \mathbb{E}(f \mid [x]_i) - \mathbb{E}(f \mid [y]_i) \\ &= \mathbb{E}(f \mid [x]_i) - \mathbb{E}(f \circ \phi \mid [x]_i) \\ &= \mathbb{E}(f - f \circ \phi \mid [x]_i) \\ &\leq \mathbb{E}(d(\,.\,, \phi(\,.\,)) \mid [x]_i) && (f \text{ is 1-Lipschitz}) \\ &\leq a_i \end{aligned}$$

This concludes the proof. □

Let $\boldsymbol{X} = (X, d, \mu)$ be a finite mm-space and $A \subseteq X$ measurable. Observe that $d_A \colon X \to \mathbb{R}$, $d_A(x) := \inf_{y \in A} d(x, y)$ is a 1-Lipschitz function. Using this we can rewrite the definition of the measure concentration function

$$\alpha_{\boldsymbol{X}}(\varepsilon) = \sup\{\mu(\{d_A \geq \varepsilon\}) \mid \mu(A) \geq \frac{1}{2}\}.$$

This gives us the desired connection.

Theorem 15. *If a finite mm-space $\boldsymbol{X} = (X, d, \mu)$ has length l, then the measure concentration function of $\boldsymbol{X}$ satisfies*

$$\alpha_{\boldsymbol{X}}(\varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{16 l^2}\right) \textit{ for all } \varepsilon > 0.$$

Proof. Let $\varepsilon > 0$ and $A \subseteq X$ be measurable with $\mu(A) \geq \frac{1}{2}$. As mentioned above d_A is 1-Lipschitz and therefore, by Lemma 14,

$$\mu(\{|d_A - \mathbb{E}(d_A)| \geq \varepsilon\}) \leq 2 \exp\left(-\frac{\varepsilon^2}{4 l^2}\right).$$

Now there are two cases to consider, the first case is the more interesting one.

If $\mathbb{E}(d_A) \leq \varepsilon$, then for any x with $d_A(x) \geq 2\varepsilon$ we know $d_A(x) \geq \varepsilon + \mathbb{E}(d_A)$ and therefore $|d_A(x) - \mathbb{E}(d_A)| \geq \varepsilon$. As a consequence

$$\mu(\{d_A \geq 2\varepsilon\}) \leq \mu(\{|d_A - \mathbb{E}(d_A)| \geq \varepsilon\}) \leq 2 \exp\left(-\frac{\varepsilon^2}{4 l^2}\right).$$

Replacing ε by $\frac{\varepsilon}{2}$ gives the desired inequality.

If $\mathbb{E}(d_A) > \varepsilon$, then $A \subseteq \{|d_A - \mathbb{E}(d_A)| \geq \varepsilon\}$. Consequently,

$$\mu(\{d_A \geq \varepsilon\}) \leq \mu(X \setminus A) \leq \frac{1}{2} \leq \mu(A) \leq \mu(\{|d_A - \mathbb{E}(d_A)| \geq \varepsilon\}) \leq 2\exp\left(-\frac{\varepsilon^2}{4l^2}\right).$$

This proves the theorem. □

Note that in the second case the upper bound is at least $\frac{1}{2}$, which means that if l is large enough then we are in the first case and the expected distance to a set with at least half measure is less then ε. Our goal is to apply Theorem to groups and as it turns out we can bound the length of a group using sequences of subgroups. Before we can write down the corollary we need to make a quick excursion to factor metrics.

Definition 16. Let (X, d) be a metric space and let $\sim$ be an equivalence relation on X. Then

$$d_\sim([x],[y]) = \inf\{d(p_1, q_1) + \cdots + d(p_n, q_n) \mid q_i \sim p_{i+1}, x \sim p_1, q_n \sim y\}$$

defines a pseudometric on $X/\sim$.

In case that X is a group with bi-invariant metric this definition simplifies.

Lemma 17. *Let G be a finite group with bi invariant metric d and H a (not necessarily normal) subgroup of G. Then the factor metric d_H on $G/H = \{gH \mid g \in G\}$ is a proper metric and satisfies $d_H(gH, g'H) = \inf\{d(g, g'h) \mid h \in H\}$.*

Proof. Let $x, y \in G$. We show that for any path $p_1, q_1, \ldots, p_n, q_n$ as in the definition there are $x \sim p$ and $q \sim y$ such that $d(p, q) \leq d(p_1, q_1) + \cdots + d(p_n, q_n)$. It suffices to show this for $n = 2$. By definition p_1, q_1, p_2, q_2 are of the form $g, g', g'h, g''$ form some $g, g', g'' \in G$ and $h \in H$. Since d is bi-invariant

$$d(gh, g'') \leq d(gh, g'h) + d(g'h, g'') = d(g, g') + d(g'h, g'').$$

Furthermore we are given that G is finite. Hence the infimum becomes a minimum and $d_H([x],[y]) = 0$ only if $[x] = [y]$. □

Equipped with this knowledge we can formulate the final statement for this section.

Corollary 18. *Let G be a finite group with a bi-invariant metric d, and let*

$$\{e\} = G_0 < G_1 < \cdots < G_n = G$$

be a chain of subgroups. Denote the diameter of G_i/G_{i-1} with respect to the factor metric by a_i. Then the measure concentration function of the mm-space (G, d, μ), where μ is the normalized counting measure, satisfies

$$\alpha_X(\varepsilon) \leq 2\exp\left(-\frac{\varepsilon^2}{16 \cdot \sum_{i=1}^n a_i^2}\right).$$

Proof. We show that the length l of (G, d, μ) is bounded by $\left(\sum_{i=1}^{n} a_i^2\right)^{\frac{1}{2}}$ and apply Theorem 15. Define the sequence of partitions $\Omega_i := \{gG_i \mid g \in G\}$

$$\begin{aligned} \{\{g\} \mid g \in G\} = \Omega_0 \succ \Omega_1 \succ \cdots \succ \Omega_n &= \{G\} \\ \{e\} = G_0 < G_1 < \cdots < G_n \ &= G. \end{aligned}$$

Take $A \in \Omega_{i+1}$ and $g, g' \in A$. Since the distance of gG_i and $g'G_i$ with respect to the factor metric is at most a_i there is an $h' \in G_i$ such that $d(g, g'h') \leq a_i$. Hence the map

$$\begin{aligned} \phi\colon gG_i &\to g'G_i \\ gh &\mapsto g'h'h \end{aligned}$$

is, by bi-invariance of d, an isomorphism of metric spaces with $d(gh, g'h'h) = d(g, g'h') \leq a_i$ for all $gh \in gG_i$. Therefore $\left(\sum_{i=1}^{n} a_i^2\right)$ is an upper bound for l^2. $\square$

Carderi and Thom used this result to show that the limit of $\mathrm{SL}_{2^n}(q)$ is extremely amenable [**?**]. We will recreate this proof in the next section.

4 The limit of $\mathrm{SL}_{2^n}(q)$ is extremely amenable

When studying matrices it is often useful to look at the corresponding linear maps of a suitable vector space. In the case of $\mathrm{SL}_n(q)$ an n dimensional $\mathbb{F}_q$ vector space V suffices. Fixing a basis $e_1, \ldots, e_n$ gives us an embedding from $\mathrm{SL}_n(q)$ into $\mathrm{Aut}(V)$. Next we will apply the methods from the previous section to show that $\operatorname{clim} \mathrm{SL}_{2^n}(q)$ is extremely amenable.

Theorem 19. *The normalized counting measure on the groups $\mathrm{SL}_n(q)$ concentrates with respect to the normalized rank-metric, i.e. for all $r > 0$*

$$\lim_{n \to \infty} \alpha_{\mathrm{SL}_n(q)}(r) = 0.$$

Florian sagt: "Boldsymbol or tuple"

Florian sagt: "proof for SL and GL te same?"

Proof. We will apply Corollary 18 to a sequence of subgroups which also shows that the length of $\mathrm{SL}_n(q)$ is bounded by $3n^{-\frac{1}{2}}$. Let $e_1, \ldots, e_n$ be a basis of an n dimensional $\mathbb{F}_q$ vector space V. Look at the sequence

$$\mathrm{SL}_0(q) < \mathrm{SL}_1(q) < \cdots < \mathrm{SL}_n(q),$$

where $\mathrm{SL}_{i-1}(q)$ becomes a subgroup of $\mathrm{SL}_i(q)$ via the embedding $g \mapsto \begin{pmatrix} g & 0 \\ 0 & 1 \end{pmatrix}$. Next we want to bound the diameter of $\mathrm{SL}_i(q) / \mathrm{SL}_{i-1}(q)$ by $\frac{3}{n}$. By Lemma 17 it suffices to show that for any $g, g' \in \mathrm{SL}_i(q)$ there is an $h \in \mathrm{SL}_{i-1}(q)$ such that $d(g, g'h) \leq \frac{3}{n}$. Since d is bi-invariant we can assume w.l.o.g. that $g' = 1_V$. Our goal is now to find a $g' \in \mathrm{SL}_i(q)$ that is the identity on e_i.

Take a closer look at ge_i. If e_i is an eigenvector of g with eigenvalue λ, then $\lambda \neq 0$ and g is of the form $\begin{pmatrix} A & 0 \\ c^\perp & \lambda \end{pmatrix}$. Define $h' := \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ and $g' := \begin{pmatrix} I_{i-2} & 0 \\ 0 & h' \end{pmatrix} \cdot g$. By construction $g' \in \mathrm{SL}_i(q)$ and it is of the form $\begin{pmatrix} A' & 0 \\ c'^\perp & 1 \end{pmatrix}$. Since $\det g' = 1$ we have that $\det A' = 1$ and therefore $A' \in \mathrm{SL}_{i-1}(q)$ making it a suitable candidate for h. Using the triangle inequality we obtain

$$\begin{aligned} d(g,h) &\leq d(g,g') + d(g',h) \\ &= d(I_i, \begin{pmatrix} I_{i-2} & 0 \\ 0 & h' \end{pmatrix}) + \frac{1}{n}\mathrm{r}(\begin{pmatrix} 0 & 0 \\ -c'^\perp & 0 \end{pmatrix}) \\ &\leq \frac{2}{n} + \frac{1}{n} \end{aligned}$$

as desired.

If e_i is not an eigenvector of g, then we can make a change of basis of $\langle e_1, \dots, e_{i-1} \rangle$ such that $ge_i = e_{i-1} + \lambda e_i$. Henceforth we can assume w.l.o.g. that g is of the form $\begin{pmatrix} & A & & 0 \\ & & & 1 \\ c^\perp & & c_{i-1} & \lambda \end{pmatrix}$. Define $h' := \begin{pmatrix} \lambda & -1 \\ 1 & 0 \end{pmatrix}$ and as before $g' := \begin{pmatrix} I_{i-2} & 0 \\ 0 & h' \end{pmatrix} \cdot g$. Now we can apply the argument from above to get an $h \in \mathrm{SL}_{i-1}(q)$ such that $d(g,h) \leq \frac{3}{n}$. Applying Corollary 18 we obtain

$$\alpha_{\mathrm{SL}_n(q)}(\varepsilon) \leq 2\exp\left(-\frac{\varepsilon^2}{16 \cdot \sum_{i=1}^n \frac{9}{n^2}}\right) = 2\exp\left(-\frac{\varepsilon^2 n}{16 \cdot 9}\right),$$

which tends to 0 as n goes to infinity.

□

From this theorem the main result of this section easily follows.

Corollary 20. *The Polish group* $\operatorname{clim}\mathrm{SL}_{2^n}(q)$ *is extremely amenable.*

Proof. Theorem 19 implies that $\operatorname{clim}\mathrm{SL}_{2^n}(q)$ is a Levy group and is therefore extremely amenable, by Theorem 4. □

As a byproduct we found an upper bound for the length of $\mathrm{SL}_n(q)$. We now ask how good this upper bound is. Therefore we our next goal is to determining also a lower bound. This part is not essential to the rest of the thesis but still interesting.

Lemma 21. *Let* (X, d, μ) *be a finite mm-space with diameter* Δ *and*

$$\Omega_0 = \{X\} \prec \cdots \prec \Omega_n = \{\{x\} \mid x \in X\}$$

with $a_1, \dots, a_n$ *as in Definition 11. Then*

$$\sum_{i=1}^n a_i \geq \Delta.$$

Proof. Let $x, y \in X$, with $x \neq y$, we show $d(x,y) \leq \sum_{i=1}^{n} a_i$. Let i_0 be the smallest number such that $[x]_{i_0} \neq [y]_{i_0}$. Since $[x]_0 = X = [y]_0$ we know that i_0 is at least 1. Therefore $[x]_{i_0-1} = [y]_{i_0-1}$ and there is an isomorphism $\varphi_{i_0} \colon [x]_{i_0} \to [y]_{i_0}$ such that $d(\varphi_{i_0}(x), y) \leq a_{i_0}$. Let $x_{i_0} = \varphi_{i_0}(x)$, then

$$d(x,y) \leq d(x, x_{i_0}) + d(x_{i_0}, y).$$

If $x_{i_0} = y$, then we are done. Otherwise let i_1 be the smallest number such that $[x_{i_0}]_{i_1} \neq [y]_{i_1}$. Then let $\varphi_{i_1} \colon [x_{i_0}]_{i_1} \to [y]_{i_1}$ be an isomorphism such that $d(\varphi_{i_1}(x_{i_0}), y) \leq a_{i_1}$. Define $x_{i_1} = \varphi_{i_1}(x_{i_0})$. Proceeding in this fashion yields elements $x_{i_0}, \ldots, x_{i_k}$ such that $x_{i_k} = y$ and

$$d(x,y) \leq d(x, x_{i_0}) + d(x_{i_0}, x_{i_1}) + \cdots + d(x_{i_{k-1}}, x_{i_k}) \leq a_{i_0} + \cdots + a_{i_k} \leq \sum_{i=1}^{n} a_i.$$

From this the claim immediately follows. □

Lemma 22. *Let (X, d, μ) be a finite mm-space with diameter Δ and $\delta = \min_{x \neq x} d(x,y)$. Then the length of X is at least $(\Delta \cdot \delta)^{\frac{1}{2}}$.*

Proof. We show by induction on n that for any $0 \leq a_1 \leq \cdots \leq a_n$ with $\delta = \min_{1 \leq i \leq n} a_i = a_1$ we have

$$\sum_{i=1}^{n} a_i \geq \Delta \implies \sum_{i=1}^{n} a_i^2 \geq \Delta \cdot \delta.$$

□

Using this we can give an interval for the length of $\mathrm{SL}_n(q)$.

Corollary 23. *Consider $(\mathrm{SL}_n(q), d, \mu)$, where d is the normalized rank-metric and μ is the normalized counting measure. Then the length l of this mm-space satisfies*

$$n^{-\frac{1}{2}} \leq l \leq 3n^{-\frac{1}{2}}.$$

Proof. The diameter of $\mathrm{SL}_n(q)$ is equal to 1 and for any $g \neq g' \in \mathrm{SL}_n(q)$ we have $d(g, g') \geq \frac{1}{n}$. □

The next goal is to show that limits of symplectic groups are also extremely amenable. Theses groups can be seen as automorphism groups of a vector space together with a symplectic form. The proof will be similar to the one for the special linear groups but extending the partial inverse h' becomes much harder. This is why in the next section we will prove Witt's Lemma which does exactly what we need, i.e. extending isometries.

Not sure whether this is still needed

Definition 24. The *symplectic group* of degree $2n$ over a field q, denoted by $\mathrm{Sp}(2n, q)$, is the subgroup of $\mathrm{SL}(2n, q)$ containing all matrices A such that

$$A^T \Omega A = \Omega, \text{ where } \Omega = \begin{pmatrix} 0 & E_n \\ -E_n & 0 \end{pmatrix}.$$

Lemma 25. *Let $g \colon V \to V$ be an isomorphism, $V = U \oplus U'$, and $g(U') \subseteq U'$. Then the map*

$$\begin{aligned} g' \colon V &\to V \\ v &\mapsto \begin{cases} g(v) - \pi_{U'}(g(v)) & \text{if } v \notin U' \\ v & \text{if } v \in U' \end{cases} \end{aligned}$$

i.e. $g' = \pi_U \circ g - \pi_U \circ g \circ 1_{U'} + 1_{U'}$, is an isomorphism and $d(g, g') \leq \frac{1}{n} \cdot \dim U'$.

Lemma 26. *[what we still need (add conditions for ω if necessary)] Let $\omega \colon V \times V \to k$ be a bilinear map, U, U' subspaces of V, and $h \colon U \to U'$ an isomorphism that preserves ω. Then h can be extended to an isomorphism on V which also preserves ω.*

Proof. w.l.o.g. $\dim U + 1 = \dim V$? □

Florian sagt: "…additional conditions"

Lemma 27. *Let $V = U \oplus U'$, ω a bilinear map, G be the group of automorphisms of (V, ω) and $G' \leq G$ the subgroup fixing U'. Then the diameter of G/G_i is at most $\frac{3 \cdot \dim U'}{n}$.*

Florian sagt: "adapt this"

Proof. Let $g \in G$, we show that there are $g' \in G$ and $g'' \in G'$ such that $g'(U') \subseteq U'$, $g'|_{U'} = 1_{U'}$, and

$$d(g, g'') \leq d(g, g') + d(g', g'') \leq \frac{2 \dim U'}{n} + \frac{\dim U'}{n}.$$

By Lemma 26 we can extend the map $g^{-1}|_{gU'}$ to a map h' on $V' = \langle U', gU' \rangle$. Now define $g' = (1_{V''} \oplus h')g$, where $V = V'' \oplus V'$ and apply Lemma to g' to obtain g''.

$$\begin{aligned} \operatorname{im} g - g' &= \operatorname{im} g - (1_{V''} \oplus h')g \\ &= \operatorname{im}(1_{V''} \oplus 1_{V'} - 1_{V''} \oplus h') \\ &= \operatorname{im}(1_{V'} - h') \\ &\subseteq V' \end{aligned}$$

$$d(g, g') = \frac{1}{n} \dim \operatorname{im} g - g' \leq \frac{\dim V'}{n}$$

□

5 Witts Lemma

6 Limits of other Matrix group families are Levy groups too

When studying matrices it is often useful to look at the corresponding linear maps of a suitable vector space. In the case of orthogonal, symplectic, or unitary matrices these are linear maps from the vector space to itself preserving an orthogonal, symplectic, or unitary form respectively. Formally, the symplectic group $\mathrm{Sp}_n(q)$ is isomorphic to $\mathrm{Aut}(V, \omega)$, where V is an n-dimensional $F(q)$ vector space and ω is a symplectic form.

As we have to handle only finite dimensional vector spaces here a lot of nice theorems hold. …

Let V be an n dimensional vector space.

Lemma 28. *For all $U \leq V$ there is an $U' \leq V$ such that $U \oplus U' = V$.*

Let ω be a bilinear form on V.

Lemma 29. *Let $U \leq V$. Then $\dim U^\perp = \dim V - \dim U$.*

Lemma 30. *Let $U \leq V$. Then $U^{\perp^\perp} = U$.*

Lemma 31. *There exists a $U \leq V$ with $\dim U \leq 2$ such that $V = U \oplus U^\perp$.*

Proof. Let $e \in V \setminus \{0\}$. By Lemma 29 $\dim e^\perp = n - 1$.

If $e \notin e^\perp$, then $V = \langle e \rangle \oplus e^\perp$ and $\langle e \rangle$ is the desired U.

If $e \in e^\perp$, then extend e to a basis $e, b_2, \ldots, b_{n-1}$ of $e^\perp$ and consider the 2-dimensional subspace $U := \langle b_2, \ldots, b_{n-1} \rangle^\perp$. Now we have to show that

$$U \cap U^\perp = 0.$$

Take v from the intersection. By Lemma 30 $U^\perp = \langle b_2, \ldots, b_{n-1} \rangle$ and $v \perp b_i$ for all $i \in \{2, \ldots, n-1\}$. Since $\langle b_2, \ldots, b_{n-1} \rangle \leq e^\perp$ we also have $v \perp e$. Hence $v \in e^{\perp^\perp} = \langle e \rangle$ and $v = \lambda e$. Now $e \notin \langle b_2, \ldots, b_{n-1} \rangle$ implies $v = 0$. Henceforth $V = U \oplus U^\perp$. □

Lemma 32. *Let $U \leq V$ and $f \colon V \to V$ be an isometry such that $f|_U = 1_U$. Then $f(U^\perp) = U^\perp$.*

Proof. As $\dim f(U^\perp) = \dim U^\perp$ it suffices to show that $f(u') \perp u$ for all $u \in U$ and $u' \in U^\perp$.

$$\begin{aligned} \omega(f(u'), u) &= \omega(f(u'), f(u)) \\ &= \omega(u', u) \\ &= 0 \end{aligned}$$

This concludes the proof. □

Lemma 33. *For all $W \leq V$ there is a $W' \leq W^{\perp}$ such that $W \cap W' = 0$ and*

$$\dim W' \geq \dim V - 2 \dim W.$$

Proof. By Lemma 28 there is a W' such that

$$W^{\perp} = (W^{\perp} \cap W) \oplus W'.$$

Clearly, $W \cap W' = 0$ and

$$\dim W' = \dim W^{\perp} - \dim(W^{\perp} \cap W) \geq \dim W^{\perp} - \dim W.$$

Whats left is to show that $\dim W^{\perp} \geq \dim V - \dim W$. Let $b_1, \ldots, b_{\dim W}$ be a basis of W. Then $W^{\perp}$ is equal to the kernel of the linear map

$$V \to F_q^{\dim W} \qquad v \mapsto \begin{pmatrix} \omega(b_1, v) \\ \vdots \\ \omega(b_{\dim W}, v) \end{pmatrix}.$$

Now the statement follows from the rank-nullity theorem. $\square$

Lemma 34. *Let $U, W \leq V$ such that $U \perp W$ and $U \cap W = 0$. Then $\langle U, W \rangle \cong U \oplus W$.*

Lemma 35. *Let $g_1 \colon U_1 \to W_1$ and $g_2 \colon U_2 \to W_2$ be isometries such that $U_1 \perp U_2$, $U_1 \cap U_2 = 0$, $W_1 \perp W_2$, and $W_1 \cap W_2 = 0$. Then $g_1 \oplus g_2 \colon U_1 \oplus U_2 \to W_1 \oplus W_2$ is also an isomtry.*

Florian sagt: "maybe $g \colon U_1 \to U_2$ and $h \colon W_1 \to W_2$ better"

Proof. Obviously, $g_1 \oplus g_2$ is again a bijective linear map. Consider $v_1 + v_2, u_1 + u_2 \in U_1 \oplus U_2$

$$\begin{aligned}
\omega(v_1 + v_2, u_1 + u_2) &= \omega(v_1, u_1) + \omega(v_1, u_2) + \omega(v_2, u_1) + \omega(v_2, u_2) \\
&= \omega(v_1, u_1) + 0 + 0 + \omega(v_2, u_2) && (U_1 \perp U_2) \\
&= \omega(g_1(v_1), g_1(u_1)) + \omega(g_2(v_2), g_2(u_2)) \\
&= \omega(g_1(v_1), g_1(u_1)) + \omega(g_1(v_1), g_2(u_2)) \\
&\quad + \omega(g_2(v_2), g_1(u_1)) + \omega(g_2(v_2), g_2(u_2)) && (W_1 \perp W_2) \\
&= \omega(g_1 \oplus g_2(v_1 + v_2), g_1 \oplus g_2(u_1 + u_2))
\end{aligned}$$

Hence $g_1 \oplus g_2$ preserves ω. $\square$

[other useful theorems]

Theorem 36 (Witt)**.** *Let V be an orthogonal, symplectic, or unitary space. Let U and W be subspaces of V and suppose $\alpha \colon U \to W$ is an isometry. Then α extends to an isometry of V.*

Lemma 37. *Let G be an orthogonal, symplectic, or unitary group. ...*

Proof. $G = \mathrm{Aut}(V, \omega)$ for some vector space V with bilinear form ω. Use Lemma 31 to obtain $U \leq V$ such that $V = U \oplus U^{\perp}$ and $\dim U \leq 2$. Define $H = \mathrm{Aut}(U^{\perp}, \omega)$. Our aim is to find for any $g \in G$ an $g' \in H$ such that $d(g, g') \leq \frac{8}{n}$. The idea is to find a map $h \in H$ that behaves like the inverse of g on gU and like the identity on most of the rest. Then hg is the desired g'.

Let $g \in G$ and define $W = \langle U, gU \rangle$. By Lemma 33 there is a W' such that $\dim W' \geq n - 8$, $W' \leq W^{\perp}$, and $W' \cap W = 0$. Consider the map

$$g^{-1}|_{gU} \oplus 1_{W'} \colon gU \oplus W' \to U \oplus W'$$

as $g^{-1}|_{gU}$ and $1_{W'}$ are isometries and $W \perp W'$ Lemma 35 implies that the above map is also an isomtry. By Witt's lemma this isometry can be extended to an isometry $h \colon V \to V$.

$$\begin{aligned}
n \cdot d(g, hg) &= \dim \mathrm{im}(g - hg) \\
&\leq 8 + \dim \mathrm{im}(g - hg)|_{W'} && (\dim W' \geq n - 8) \\
&= 8 + \dim \mathrm{im}(g - g)|_{W'} && (h|_{W'} = 1_{W'}) \\
&= 8
\end{aligned}$$

Finally, we need to show that $hg \in H$, here the choice of H using Lemma 31 comes into play. By construction of h we have that $hg|_U = 1_U$. Therefore we can apply Lemma 32 and get that $hg(U^{\perp}) = U^{\perp}$. Hence $hg \in H$ and $d(g, hg) \leq \frac{8}{n}$. $\square$

7 Fun

Consider an n-dimensional cube with 2^k nodes on each edge. Then its diameter $\nabla_{n,k}$ and length $L_{n,k}$ are

$$\nabla_{n,k} = \sqrt{(2^k - 1) \cdot n} \qquad\qquad L_{n,k} = \sqrt{\sum_{i=0}^{k-1} 2^{2i} \cdot n}.$$

Henceforth

$$\lim_{n \to \infty} \frac{L_{n,k}}{\nabla_{n,k}} = \frac{L_{1,k}}{\nabla_{1,k}} \qquad \text{and} \qquad \lim_{k \to \infty} \frac{L_{n,k}}{\nabla_{n,k}} = \frac{1}{\sqrt{3}}.$$

ERKLÄRUNG

Hiermit erkläre ich, dass ich die am heutigen Tag eingereichte Diplomarbeit zum Thema "Topological Entropy of Formal Languages" selbstständig erarbeitet, verfasst und Zitate kenntlich gemacht habe. Andere als die angegebenen Hilfsmittel wurden von mir nicht benutzt.

Datum Unterschrift