# Non-archimedean Excursions in Probability, Number theory, Combinatorics and Geometry

by

Yassine El Maazouz

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Statistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bernd Sturmfels, Chair
Professor Steven Neil Evans,
Professor Jim Pitman,
Professor Kiran Kedlaya

Fall 2022

Non-archimedean Excursions in
Probability, Number theory, Combinatorics and Geometry

Abstract

Non-archimedean Excursions in
Probability, Number theory, Combinatorics and Geometry

by

Yassine El Maazouz

Doctor of Philosophy in Statistics

University of California, Berkeley

Professor Bernd Sturmfels, Chair

Similar to the field of real numbers $\mathbb{R}$, which can be constructed as the completion of the rational numbers with respect the usual absolute value $|\cdot|$, the field of $p$-adic numbers $\mathbb{Q}_p$ is also the completion of the rational numbers with the $p$-adic absolute value $|\cdot|_p$. These numbers were first introduced by Kurt Hensel to harness the power of analytic tools in number theory, but nowadays non-archimedean mathematics has become a very rich and active area of research in its own right. This thesis is composed of seven chapters whose main theme is non-archimedean.

This dissertation begins with the first chapter in which we introduce some necessary background and concepts that are used throughout this thesis. In particular, we recall the notion of valued fields and review some basic facts from non-archimedean algebra and analysis. We then introduce the Euclidean building associated to the reductive group PGL, which will appear throughout this thesis.

In the second chapter we study the entropy map for multivariate Gaussian distributions on a non-archimedean local field. As in the real case, the image of this map lies in the supermodular cone. Moreover, given a multivariate Gaussian measure on a local field, its image under the entropy map determines its pushforward under the valuation map. In general, this entropy map can be defined for non-archimedian valued fields whose valuation group is an additive subgroup of the real line, and it remains supermodular. We also explicitly compute the image of this map in dimension 3 and discuss non-archimedean statistical Gaussian models.

In the third chapter we give a method for sampling points from an algebraic manifold (affine or projective) over a local field with a prescribed probability distribution. In the spirit of previous work by Breiding and Marigliano on real algebraic manifolds, our method is based on slicing the given variety with random linear spaces of complementary dimension. We also

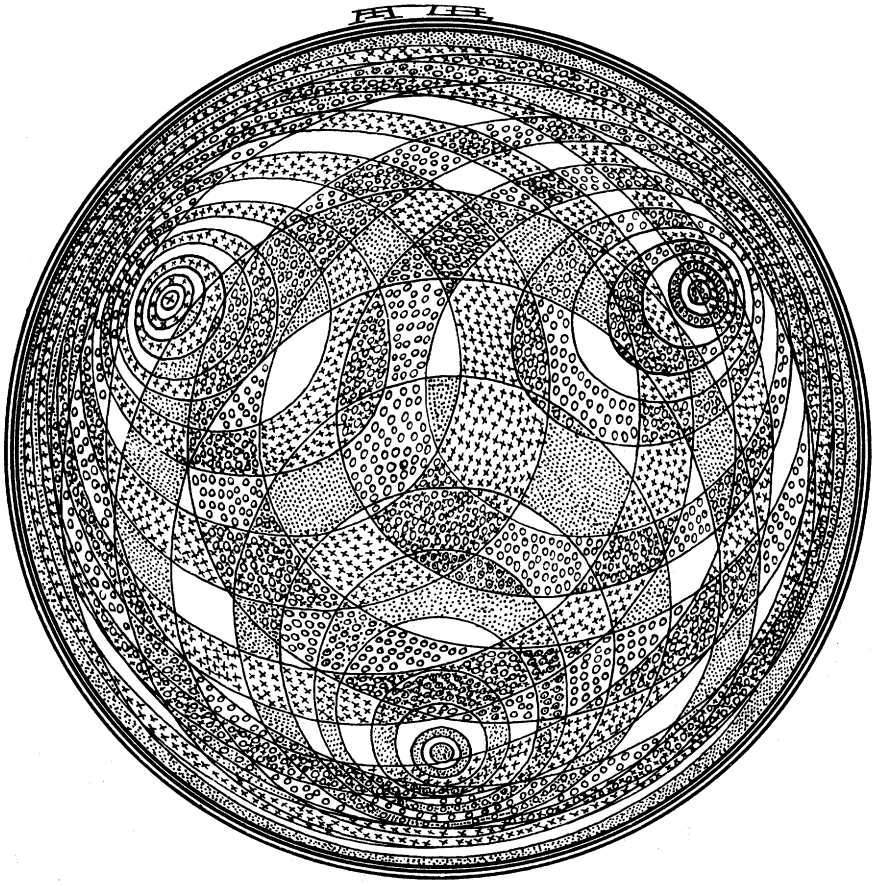
provide an implementation of our sampling method and discuss a few applications. In one application, we sample from algebraic $p$-adic matrix groups and modular curves.

In the fourth chapter, we introduce a novel characterization of Bernoulli polynomials by circular convolution. There is a combinatorial and probabilistic model underlying this characterization which we call *The Bernoulli clock*. We use this model to give a probabilistic perspective to the work of Horton and Kurn [91] and the more recent work of Clifton et al. [33] on counting permutations of the multiset $1^m \cdots n^m$ with longest continuous and increasing subsequence of length $n$ starting from 1.

In the fifth chapter we apply tropical geometry to study matrix algebras over a field with valuation. Using the shapes of min-max convexity, known as polytropes, we revisit the graduated orders introduced by Plesken and Zassenhaus. These are classified by the polytrope region. We also advance the ideal theory of graduated orders by introducing their ideal class polytropes. We then extend our study to *bolytropes* and *bolytrope orders*. Bolytropes are bounded subsets of an affine building that consist of all points that have distance at most $r$ from some polytrope. We prove that the points of a bolytrope describe the set of all invariant lattices of a bolytrope order, generalizing the correspondence between polytropes and graduated orders.

In the sixth chapter, we study non-archimedean Schur representations and their invariant lattices. More precisely, fix a non-archimedean discretely valued field $K$ and let $\mathcal{O}_K$ be its valuation ring. Given a vector space $V$ of dimension $n$ over $K$ and a partition $\lambda$ of an integer $d$, we study the problem of determining the invariant lattices in the Schur module $S_\lambda(V)$ under the action of the group $\mathrm{GL}(n, \mathcal{O}_K)$. When $K$ is a non-archimedean local field, our results determine the $\mathrm{GL}(n, \mathcal{O}_K)$-invariant Gaussian distributions on $S_\lambda(V)$.

Finally, in the seventh chapter, we turn to arithmetic geometry. We express the reduction types of Picard curves in terms of tropical invariants associated to binary quintics. These invariants are connected to Picard modular forms using recent work [32] of Cléry and van der Geer. We furthermore give a general framework for tropical invariants associated to group actions on arbitrary varieties. The problem of describing reduction types of curves in terms of their associated invariants fits in this general framework by mapping the space of binary forms to symmetrized versions of the Deligne–Mumford compactification $\overline{M}_{0,n}$.



Artist's conception of the 3-adic unit disk.

*Drawing by A.T. Fomenko of Moscow State University, Moscow, U.S.S.R.*

*To my father.*

# Contents

# List of Figures

# List of Tables

## Acknowledgments

I owe a great deal of recognition to many people without whom I would not have written this dissertation. It is my duty and pleasure to express my deep feelings of gratitude to anyone who gave me a push along the way, and I apologize in advance to anyone I, somehow, forgot to mention.

First, I would like to thank my advisor, Bernd Sturmfels. Bernd, thank you for your wise guidance, unwavering support and infinite energy. You have been far more than an advisor to me and I am enormously grateful for the opportunity to learn from you, both mathematically and otherwise. Next, I want to thank my faculty mentor, Jim Pitman. Jim, many thanks for your constant enthusiasm, encouragement and kind spirit. I learned a lot during our numerous discussions, hikes and walks (not to mention a very memorable tennis lesson). I count myself lucky to have gotten the chance to discuss with you as much as I did. A very special thanks to my second faculty mentor in Berkeley, Steve Evans, who always had time and patience to answer my questions with great precision and whose work inspired a large portion of this thesis. I was very fortunate to have gotten the opportunity to learn from you. I also thank Kiran Kedlaya whom I only met virtually so far, yet had the kindness to be on my qualifying exam and dissertation committees and was extremely generous with his time whenever I had a question.

The research presented in this thesis is mainly based on joint work with my collaborators: Enis Kaya, Paul A. Helminck, Marvin A. Hahn, Gabriele Nebe, Antonio Lerario, Jim Pitman, Mima Stanojkovski, Bernd Sturmfels and Ngoc Tran. Thank you, I learned so much from working with you.

Albeit shorter than I expected, my time at Berkeley was quite the ride: exciting, challenging, unusual and, most important, very instructive. I had the opportunity to interact with and learn from some of the most talented and brilliant mathematicians. Equally important, I have gotten to know many awe-inspiring and extremely kind people and made a few amazing friends who made Berkeley my home for the last three years. In particular, I would like to thank Alice and Corrine for their extreme kindness, friendship and support. Many thanks to the members of the statistics department I had the chance to interact with. In particular, I thank La Shana Porlaris for her amazing work (especially getting me through my qualifying exam in the nick of time). Many thanks also to my academic siblings in Berkeley: Yelena and Yulia.

A considerable amount of my research was done during my frequent visits to the Max Planck institute for mathematics in the sciences in Leipzig. I thank Bernd for the generous invitations, Saskia for her amazing work to ensure my visits to Leipzig were as smooth as can be and the institute for the warm hospitality. I also thank my academic siblings in Leipzig: Chiara, Claudia, Kemal and Rida, and all the members of the *Non-linear algebra group* I had the chance to meet. Last but not least, I thank my family for their continuous support and patience.

# Chapter 1

# Background and introduction

This chapter is meant to collect some basic notions and background in non-archimedean algebra and analysis, and briefly introcude the Bruhat-Tits building associated to the reductive group PGL which will be used repeatedly in the remaining chapters. Unless it is necessary for our later discussions, we shall not give proofs for any stated result. Instead we shall refer the reader to more complete and authoritative sources whenever necessary.

## 1.1 Non-archimedean algebra and analysis

The material in this section is well known and can be found for example in [57, 144, 146]

### 1.1.1 Valued fields

Let us start by introducing the notion of a (non-archimedean) valued field, which is a central object in our study.

**Definition 1.1.** A valued field $(K, |\cdot|)$ is a field $K$ together with a map $|\cdot|\colon K \to \mathbb{R}_+$ satisfying the following three conditions for any $x, y \in K$:

$$\begin{aligned} &(i)\ |x| = 0 \text{ if and only if } x = 0, && \text{(separation)},\\ &(ii)\ |xy| = |x||y|, && \text{(multiplicativity)},\\ &(iii)\ |x+y| \leq |x| + |y| && \text{(triangle inequality)}. \end{aligned}$$

The map $|\cdot|$ is called an *absolute value*. It is also referred to as a *norm* or *multiplicative valuation* on $K$.

**Example 1.2.** The field $\mathbb{R}$ of real numbers together with its usual absolute value is a valued field, as is the field $\mathbb{C}$ with the usual modulus map $|\cdot|$.

**Definition 1.3.** Let $(K, |\cdot|)$ be a valued field. A non-zero element $x \in K$ is called archimedean if the set $\{|nx| = |x+\ldots+x|\colon n \in \mathbb{N}\}$ is unbounded in $\mathbb{R}_+$, and non-archimedean otherwise.

Notice that, thanks to the multiplicativity of the absolute value, a non-zero $x \in K$ is archimedean if and only if 1 is archimedean. When $1 \in K$ is archimedean, the field $K$ is called an *archimedean valued field*, and it is called *non-archimedean* otherwise.

**Example 1.4.** 1. Let $K$ be any field and let $|\cdot|_{\text{triv}}$ be the trivial norm on $K$ i.e. $|x|_{\text{triv}} = 1$ for any non-zero $x \in K$. The valued field $(K, |\cdot|_{\text{triv}})$ is non-archimedean.

2. The fields $\mathbb{R}, \mathbb{C}$ endowed with their usual absolute value are archimedean fields.

**Proposition 1.5.** *Let $(K, |\cdot|)$ be a valued field. $K$ is non-archimedean if and only if the absolute value is* ultrametric*; that is*

$$|x + y| \leq \max(|x|, |y|), \quad \textit{for any} \quad x, y \in K. \tag{1.1}$$

If $(K, |\cdot|)$ is a valued field, the absolute value $|\cdot|$ endows $K$ with the structure of a metric space with the distance given by

$$d(x, y) = |x - y| \text{ for } x, y \in K,$$

and hence also with the corresponding topology (for which $K$ is a topological field).

**Proposition 1.6.** *Let $(K, |\cdot|)$ be a non-archimedean valued field. The map* $\text{val} : K^\times \to \mathbb{R}$ *given by* $\text{val}(x) = -\log(|x|)$ *is a group homomorphism. Moreover, with the convention* $\text{val}(0) = +\infty$*, we have:*

$$\text{val}(x + y) \geq \min(\text{val}(x), \text{val}(y)), \quad \textit{for } x, y \in K.$$

When the valued field $(K, |\cdot|)$ is non-archimedean, the map val is called an *additive valuation* on $K$. In this case the unit ball

$$\begin{aligned} \mathcal{O}_K &\coloneqq \{x \in K : |x| \leq 1\} \\ &= \{x \in K : \text{ val}(x) \geq 0\}, \end{aligned}$$

is a local subring of $K$; that is $\mathcal{O}_K$ has a unique maximal ideal which is

$$\begin{aligned} \mathfrak{m}_K &\coloneqq \{x \in K : |x| < 1\} \\ &= \{x \in K : \text{ val}(x) > 0\}. \end{aligned}$$

The ring $\mathcal{O}_K$ is called the *valuation ring* of $(K, |\cdot|)$ and the group of units of $\mathcal{O}_K$ is the unit circle

$$\begin{aligned} \mathcal{O}_K^\times &\coloneqq \{x \in K : |x| = 1\} \\ &= \{x \in K : \text{ val}(x) = 0\}. \end{aligned}$$

The ideal $\mathfrak{m}_K$ being maximal, the quotient $k \coloneqq \mathcal{O}_K/\mathfrak{m}_K$ is a field which is called the *residue field* of $K$.

**Remark 1.7.** The field of fractions of the ring $\mathcal{O}_K$ is exactly the field $K$.

It is known that the completion of $K$ (as a metric space) is a valued field which we denote by $\widehat{K}$ endowed with the absolute value

$$|[(x_n)]| \coloneqq \lim_{n\to\infty} |x_n|,$$

for any equivalence class $[(x_n)] \in \widehat{K}$ of Cauchy sequences in $K$. The valuation ring of the field $\widehat{K}$ is the completion $\widehat{\mathcal{O}}_K$ of $\mathcal{O}_K$ and the residue field of $\widehat{K}$ is isomorphic to the residue field of $K$.

**Example 1.8.** (The $p$-adic numbers) In this example we introduce the fundamental prototype of a *local field* which will appear numerous times in the next chapters.

Let $p$ be a prime number and let us define an absolute $|\cdot|_p$ value on $\mathbb{Q}$ as follows: for any non-zero rational $r \in \mathbb{Q}$, write $r = p^n r'$ with $n \in \mathbb{Z}$ and $r' = a'/b'$ with $a', b'$ are coprime integers that are not divisible by $p$. We then define the $p$-adic absolute value of $r$ as

$$|r|_p = p^{-n}.$$

Together with $|0|_p = 0$, this defines an absolute value on $\mathbb{Q}$ and $(\mathbb{Q}, |\cdot|_p)$ is a non-archimedean valued field with an additive valuation

$$\mathrm{val}_p \colon \mathbb{Q} \to \mathbb{Z} \cup \{+\infty\}, \quad r \mapsto -\log(|r|_p).$$

The valuation ring of $(\mathbb{Q}, |\cdot|_p)$ is the local ring

$$\mathbb{Z}_{(p)} = \left\{ \frac{a}{b} \colon a, b \in \mathbb{Z} \text{ and } b \text{ is not divisible by } p \right\},$$

its unique maximal ideal is $p\mathbb{Z}_{(p)}$ and the residue field is the finite field with $p$ elements:

$$\mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)} = \mathbb{Z}/p\mathbb{Z}.$$

The field $\mathbb{Q}$ is not complete with respect to $|\cdot|_p$. Its completion, which we usually denote by $\mathbb{Q}_p$, is called the field of *$p$-adic numbers*. Any element $x \in \mathbb{Q}_p$ can be written uniquely in the form

$$x = \sum_{n=\mathrm{val}_p(x)}^{+\infty} a_n p^n,$$

with $a_n \in \{0, 1, \ldots, p-1\}$ for $n \geq \mathrm{val}_p(x)$. The valuation ring of $\mathbb{Q}_p$ is the ring of $p$-adic integers

$$\mathbb{Z}_p = \left\{ \sum_{n=0}^{+\infty} a_n p^n \colon 0 \leq a_n \leq p-1 \right\} = \lim_{\leftarrow n} \mathbb{Z}/p^n\mathbb{Z},$$

and its residue field is $\mathbb{Z}_p/p\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$.

**Remark 1.9.** We say that two absolute values on a field $K$ are equivalent if they define the same metric topology on $K$. It turns out that, if $|\cdot|$ is an absolute value on $\mathbb{Q}$, it is then equivalent to either the usual[1] absolute value $|\cdot|_\infty$, or to a $p$-adic absolute value $|\cdot|_p$ for some prime $p$.

### 1.1.2 Local fields

Local fields are a very important class of valued fields. They arise quite naturally in number theory as completions of number fields with respect to some valuation. In this section, we shall only see a few basic concepts which we will need in later chapters. There is an extensive literature on local fields in number theory [29], analysis [140, 144], representation theory [36].

**Definition 1.10.** A valued field $(K, |\cdot|)$ is called a local field if it is locally compact, non-discrete as a topological field.

Let $(K, |\cdot|)$ be a local field. Since $K$ is a locally compact topological field, there exists a unique Haar measure $\mu$ on $K$ with $\mu(\mathcal{O}_K) = 1$; that is a measure $\mu$ (defined on Borel subset of $K$ as a metric space) such that

$$\mu(\mathcal{O}_K) = 1 \text{ and } \mu(x + A) = \mu(A) \text{ for any } x \in K \text{ and any Borel set } A \subset K.$$

**Remark 1.11.** Let $x \in K$ and let us define the measure $\nu_x$ on $K$ as follows

$$\nu_x(A) = \mu(xA), \quad \text{for any Borel set } A.$$

The measure $\nu_x$ is also a Haar measure on $K$ so[2] there exists $\Delta(x) \geq 0$ such that

$$\nu_x = \Delta(x)\mu.$$

The map $\Delta\colon K \to \mathbb{R}_+$ is called the modulus of $K$. It turns out that $\Delta = |\cdot|$ is the absolute value in $K$ i.e. we can recover the absolute value $|\cdot|$ as follows:

$$|x| = \mu(x\mathcal{O}_K), \quad \text{for } x \in K.$$

**Example 1.12.**

1. The fields of real and complex numbers $\mathbb{R}$ and $\mathbb{C}$ are local fields.
2. The field of $p$-adic numbers is also a local field since $\mathbb{Z}_p$ is a compact and open set in $\mathbb{Q}_p$.

[1] The notation $|\cdot|_\infty$ signifies that the usual absolute value on $\mathbb{Q}$ corresponds to "a prime at infinity".
[2] Since the Haar measure on a an abelian topological is unique up to scaling.

3. Let $\mathbb{F}_p((\varpi))$ be the field of Laurent series in one variables $\varpi$ with coefficients in the finite field with $p$ elements $\mathbb{F}_p$. Elements of this field are formal series of the form

$$x = \sum_{n=v}^{+\infty} a_n \varpi^n, \quad \text{with } v \in \mathbb{Z} \text{ and } a_n \in \mathbb{F}_p \text{ for } n \geq v.$$

The valuation and absolute value of such an element can be defined as

$$\operatorname{val}(x) = \inf\{n \geq v \colon a_n \neq 0\} \quad \text{and } |x| = p^{-\operatorname{val}(x)},$$

and the valuation ring is the ring of power series

$$\mathbb{F}_p[[\varpi]] = \lim_{\leftarrow n} \mathbb{F}_p[\varpi]/\varpi^n \mathbb{F}_p[\varpi].$$

Endowed with this valuation, $\mathbb{F}_p((\varpi))$ is a local field of positive characteristic.

It turns out that the fields listed in Example 1.12 are essentially all the possible local fields in the following sense.

**Theorem 1.13.** *Let $(K, |\cdot|)$ be a local field. Then the following holds*

*(i) If $K$ is archimedean, then $K$ is isomorphic (algebraically and analytically) to $\mathbb{R}$ or $\mathbb{C}$.*

*(ii) If $K$ is non-archimedean and $\operatorname{char}(K) = 0$, then $K$ is isomorphic (algebraically and analytically) to a finite field extension of $\mathbb{Q}_p$.*

*(iii) If $K$ is non-archimedean and $\operatorname{char}(K) > 0$, then $K$ is isomorphic (algebraically and analytically) to a finite field extension of $\mathbb{F}_p((\varpi))$.*

**Remark 1.14.** Since our main focus will be non-archimedean local fields, in the remaining of this dissertation we refer to non-archimedean local fields simply as local fields.

When $K$ is a local field, the maximal ideal $\mathfrak{m}_K$ is monogenic in $\mathcal{O}_K$ (i.e. principal and generated by one element). If $\varpi \in \mathcal{O}_K$ is such a generator, that is $\mathfrak{m}_K = \varpi \mathcal{O}_K$, we call $\varpi$ a *uniformizer* of $K$. It follows from Theorem 1.13 that the residue field $k = \mathcal{O}_K/\varpi\mathcal{O}_K$ of a local field $K$ is finite.

**Proposition 1.15.** *Let $K$ be a non-archimedean local field and fix a uniformizer $\varpi$ of $K$. Let $k = \mathcal{O}_K/\varpi\mathcal{O}_K$ be the residue field and $\mathcal{T} \subset \mathcal{O}_K$ a set of representative of elements of $k$ such that $0 \in \mathcal{T}$. Then, any non-zero element $x \in K$ can be uniquely written in the form*

$$x = \sum_{n \geq v}^{+\infty} x_n \varpi^n, \quad \text{with } v \in \mathbb{Z}, \quad x_n \in \mathcal{T} \text{ and } x_v \neq 0.$$

*The valuation of such an element $x$ is then $v \in \mathbb{Z}$.*

**Remark 1.16.** Similar to the archimedean local fields $\mathbb{R}$ and $\mathbb{C}$, Fourier theory can also be carried out on any non-archimedean local field, see [159]. This theory is very fruitful in characterizing Gaussian measures over a non-archimedean local field, see [65].

### 1.1.3 Non-archimedean orthogonality

In this section, we introduce the notion of non-archimedean orthogonality. This concept will be later useful to define Gaussian measures over local fields as in [65]. The material in this section can be found with more details in [140, Chapter 5].

Let $K$ be a non-archimedean local field with normalized valuation $\mathrm{val}\colon K \to \mathbb{Z} \cup \{+\infty\}$ and fix a uniformizer $\varpi$ of $K$. Let $(E, \|\cdot\|)$ be a be a finite dimensionl vector space over $K$ and $n \coloneqq \dim_K(E)$. It is know that, since $|\cdot|$ is ultrametric, the norm $\|\cdot\|$ is also ultrametric i.e.

$$\|u+v\| \leq \max(\|u\|, \|v\|), \quad \text{for any } u, v \in E.$$

The norm $\|\cdot\|$ induces an additive valuation on $E$ defined as follows

$$\mathrm{val}(x) \coloneqq -\log(\|x\|), \quad \text{for } x \in E.$$

The valuation val satisfies the following properties

$$\begin{aligned}
&(1) \quad \mathrm{val}(x) = +\infty \text{ if and only if } x = 0, \quad \text{for all } x \in E,\\
&(2) \quad \mathrm{val}(\alpha x) = \mathrm{val}(\alpha) + \mathrm{val}(x), \quad \text{for all } \alpha \in K, x \in E,\\
&(3) \quad \mathrm{val}(x+y) \geq \min(\mathrm{val}(x), \mathrm{val}(y)), \quad \text{for all } x, y \in E.
\end{aligned}$$

Let $\Lambda$ be denote the unit ball in $E$ i.e.

$$\Lambda \coloneqq \{x \in E\colon \|x\| \leq 1\}.$$

Then $\Lambda$ is an $\mathcal{O}_K$-submodule of $E$ and $\Lambda$ spans $E$ as a vector space over $K$.

**Remark 1.17.** Let $L$ be an $\mathcal{O}_K$-submodule of $E$ such that $L$ generates $E$ over $K$ ($L$ is called a *lattice* in $E$). The lattice $L$ defines an additive valuation $\mathrm{val}_L$ on $E$ defined as follows

$$\mathrm{val}_L(x) \coloneqq \sup\left\{m \in \mathbb{Z}\colon \varpi^{-m} x \in L\right\}, \quad x \in E.$$

This valuation in turn defines a norm $\|\cdot\|_L$ on $E$

$$\|x\|_L \coloneqq q^{-\mathrm{val}_L(x)}, \quad \text{for any } x \in E.$$

We then see that there is a correspondence between lattices in $E$ and norms on $E$. We refer to [166, Chapter 2] for more details.

**Definition 1.18.** Let $(x_i)_{i\in I}$ be a family of non-zero vectors in $E$ index by some set $I$. We say that the family $(x_i)_{i\in I}$ is orthogonal in $(E, \|\cdot\|)$ if for any finite subset $J \subset I$ we have

$$\left\|\sum_{j\in J} \alpha_j x_j\right\| = \max_{j\in J} |\alpha_j| \, \|x_j\|, \quad \text{for any choice of scalars } \alpha_j \in K.$$

We say that $(x_i)_{i\in I}$ is orthonormal if it is orthogonal and $\|x_i\| = 1$ for all $i \in I$.

**Remark 1.19.** Notice that if a family $(x_i)_{i\in I}$ is orthonormal then it is necessarily linearly independent over $K$. Notice also that if a vector $x \in E$ has norm 1 then $x \in \Lambda$ and the reduction $\overline{x}$ of $x$ modulo $\varpi\Lambda$ is nonzero.

The following proposition provides a practical criterion of orthogonality.

**Proposition 1.20.** *Let $(x_i)_{i\in I} \in E$ be a family of vectors of norm 1 i.e. $\|x_i\| = 1$ for any $i \in I$. Then $(x_i)_{i\in I}$ is orthogonal if and only if the reductions $\overline{x_i}$ modulo $\varpi\Lambda$ of the $x_i$'s are linearly independent over the residue field $k$ i.e.*

$$(\overline{x_i})_{i\in I} \text{ is linearly independent over } k = \mathcal{O}_K/\varpi\mathcal{O}_K.$$

*A family $(e_1, \ldots, e_n)$ is an orthonormal basis of $E$ if and only if it is an $\mathcal{O}_K$-basis of $\Lambda$ as an $\mathcal{O}_K$-module i.e.*

$$\Lambda = \mathcal{O}_K e_1 \oplus \ldots \oplus \mathcal{O}_K e_n.$$

**Definition 1.21.** We define the orthogonal group $\mathrm{GL}(\Lambda)$ of $(E, \|\cdot\|, \Lambda)$ as follows

$$\begin{aligned}\mathrm{GL}(\Lambda) &\coloneqq \{g \in \mathrm{GL}(E)\colon g\Lambda = \Lambda\}\\ &= \{g \in \mathrm{GL}(E)\colon \|gx\| = \|x\|, \text{ for any } x \in E\}.\end{aligned}$$

**Remark 1.22.** The group $\mathrm{GL}(\Lambda)$ is a compact topological group.

Let us fix a basis $e_1, \ldots, e_n$ of $E$ and endow $E$ with the structure of a normed vector space with the lattice

$$\Lambda = \mathcal{O}_K e_1 \oplus \ldots \oplus \mathcal{O}_K e_n,$$

i.e. the corresponding norm $\|\cdot\|$ is given by

$$\|x\| = \max_{1\le i\le n} |x_i|, \quad \text{for any } x = x_1e_1 + \ldots + x_ne_n \in E.$$

Through the choice of basis $e_1, \ldots, e_n$, we identify $E$ with $K^n$ and $\Lambda$ with the lattice $\mathcal{O}_K^n$. The orthogonal group $\mathrm{GL}(\Lambda)$ in Definition 1.21 is then

$$\begin{aligned}\mathrm{GL}(\mathcal{O}_K^n) &\coloneqq \{g \in \mathrm{GL}(n, K)\colon g\mathcal{O}_K^n = \mathcal{O}_K^n\}\\ &= \{g \in \mathrm{GL}(n, K)\colon \|gx\| = \|x\| \text{ for all } x \in K^n\}\\ &= \{g \in \mathrm{GL}(n, K)\colon g \in \mathcal{O}_K^{n\times n} \text{ and } \det(g) \in \mathcal{O}_K^\times\}\\ &= \mathrm{GL}(n, \mathcal{O}_K).\end{aligned}$$

**Remark 1.23.**

1. The orthogonal group $\mathrm{GL}(n, \mathcal{O}_K)$ is also the group of matrices in $K^{n\times n}$ with orthonormal rows and columns.
2. The group $\mathrm{GL}(n, \mathcal{O}_K)$ (the orthogonal group associated to the lattice $\mathcal{O}_K^n$) is a maximal compact subgroup of $\mathrm{GL}(n, K)$. Any other maximal compact subgroup of $\mathrm{GL}(n, K)$ is of the form $\mathrm{GL}(L)$ for some lattice $L \subset K^n$ and is conjugate to $\mathrm{GL}(n, \mathcal{O}_K)$.

Let $E^\vee$ be the dual of $E$; that is $E^\vee = \mathrm{Hom}_K(E, K)$ is the vector space of linear forms on $E$. Each lattice $L$ in $E$ has a corresponding dual lattice $L^\vee := \mathrm{Hom}_{\mathcal{O}_K}(L, \mathcal{O}_K)$ in $E^\vee$. So if $(E, \|\cdot\|)$ is a normed vector space with unit ball $\Lambda$, the dual space $E^\vee$ can also be normed in a natural way by setting its unit ball to be $\Lambda^\vee$.

The following result is a non-archimedean analogue of singular value decomposition for matrices with real entries, which can be found in [62, Theorem 3.1] for example.

**Proposition 1.24** (Singular value decomposition or Smith normal form)**.** *Let $F$ be a non-archimedean discretely valued field and fix a uniformizer $\varpi$ of $F$. Let $n$ and $m$ be positive integers and let $A$ be an $n \times m$ matrix with entries in $F$. Then we can write $A$ in the form*

$$A = UDV,$$

*for some $U \in \mathrm{GL}(n, \mathcal{O}_F)$, $V \in \mathrm{GL}(m, \mathcal{O}_F)$ and $D = \mathrm{diag}(\varpi^{r_1}, \ldots, \varpi^{r_{\min(n,m)}}) \in F^{n\times m}$ such that $r_1 \geq \ldots \geq r_{\min(n,m)} \in \mathbb{Z}$.*

### 1.1.4 Gaussian measures over local fields

In this section we leverage the notion of non-archimedean orthogonality allows us to define Gaussian measures over a local field following [65]. Let $K$ be a non-archimedean local field and $(E, \|\cdot\|, \Lambda)$ a finite dimensional normed[3] vector space over $K$. We endow $E$ with the Borel $\sigma$-algebra given by $\|\cdot\|$ (i.e. the $\sigma$-algebra generated by open sets in $E$). If $\nu$ and $\mu$ are two probability measures on $E$ and $a \in K^\times$ we denote by

1. $a\mu$ the probability measure[4] on $E$ given by

$$(a\mu)(A) = \mu(a^{-1}A), \quad \text{for any borel set } A \subset E.$$

2. $\mu \otimes \nu$ the product probability measure on $E \times E$ defined by

$$(\mu \otimes \nu)(A \times B) = \mu(A)\nu(B), \quad \text{for any Borel sets } A, B \subset E.$$

3. $\mu * \nu$ the convolution of $\mu$ and $\nu$; that is the pushforward of the measure $\mu \otimes \nu$ on $E \times E$ under the addition map

$$E \times E \to E, \quad (x, y) \mapsto x + y.$$

**Definition 1.25** (Definition 4.1 of [65])**.** Let $\mathbb{P}$ be a probability distribution on $E$. We say that $\mathbb{P}$ is a Gaussian distribution if

$$\mathbb{P} \otimes \mathbb{P} = (g_{11}\mathbb{P} * g_{12}\mathbb{P}) \otimes (g_{21}\mathbb{P} * g_{22}\mathbb{P}), \quad \text{for any } g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \in \mathrm{GL}(2, \mathcal{O}_K).$$

[3] $\Lambda$ is the unit ball of $\|\cdot\|$ as in Section 1.1.3.

[4] By convention, when $a = 0$, $a\mu$ is the Dirac measure in 0 i.e. $(a\mu)(A) = 1[0 \in A]$.

Equivalently, a random variable $X$ with values in $E$ is called Gaussian if whenever $Y$ is an independent copy of $X$, the two vectors

$$\begin{pmatrix} X \\ Y \end{pmatrix} \text{ and } \begin{pmatrix} g_{11}X + g_{12}Y \\ g_{21}X + g_{22}Y \end{pmatrix}$$

have the same distribution in $E \times E$ for any $g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \in \mathrm{GL}(2, \mathcal{O}_K)$.

**Remark 1.26.** Definition 1.25 mimics Kac's characterization [98] of Gaussian distributions on Euclidean spaces (which goes back to James Clerk Maxwell's derivation of the velocity distribution for ideal gases):

> *A random variable $X$ is Gaussian in $\mathbb{R}^n$ if and only if when $Y$ is an independent copy of $X$ the following two vectors:*
>
> $$\begin{pmatrix} X \\ Y \end{pmatrix} \text{ and } \begin{pmatrix} g_{11}X + g_{12}Y \\ g_{21}X + g_{22}Y \end{pmatrix}$$
>
> *have the same distribution in $\mathbb{R}^n \times \mathbb{R}^n$ for any $g = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \in \mathrm{O}(2, \mathbb{R})$.*

Following Definition 1.25 the Fourier transform of a Gaussian distribution on $E$ should satisfy a certain invariance in the form of a functional equation. Solving this functional equation yields the following practical description of $E$-valued Gaussian measures.

**Theorem 1.27** (Theorems 4.2 and 4.4 in [65])**.** *Let $X$ be a random variable with values in $E$. Then $X$ is Gaussian if and only if the distribution of $X$ is the normalized Haar measure on an $\mathcal{O}_K$-submodule of $E$; that is there exists an $\mathcal{O}_K$-submodule $L$ of $E$ such that the distribution of $X$ is*

$$\mathbb{P}(X \in dx) = \frac{1[x \in L]}{\mu(L)} \mu(dx),$$

*where $\mu$ is any non-zero Haar measure*[5] *on $L$.*

**Remark 1.28.** In the setting of Theorem 1.27 the law of each Gaussian distribution on $E$ is completely specified by a lattice $L$. This is analogous to the setting of euclidean spaces, where the law of each centered Gaussian in $\mathbb{R}^d$ is completely specified by its covariance matrix. We note that for Gaussians over a local field the mean is not well-defined [65], so lattices in $K^d$ are indeed the central objects of the theory of Gaussian over $K$.

**Definition 1.29.** We say that a Gaussian measure on $E$ is *non-degenerate* if its corresponding submodule has full rank in $E$; i.e. is a lattice. We call *standard Gaussian distribution* on $E$ the Gaussian distribution corresponding to the unit ball $\Lambda$; that is the distribution

$$\frac{1[x \in \Lambda]}{\mu(\Lambda)} \mu(dx)$$

[5] Which exists and is unique up to scaling since $(L, +)$ is a compact topological abelian group.

for some non-zero Haar measure on $E$.

As in the Euclidean setting, independence of Gaussians is tightly linked to orthogonality as the following theorem explains.

**Theorem 1.30** (Adapted from Theorem 4.8 of [65])**.** *Let $X$ be a random variable with values in $E$ with standard Gaussian distribution. Let $f_1, \dots, f_r \in E^\vee$ be linear forms. Then the $K$-valued random variables $f_1(X), \dots, f_r(X)$ are independent if and only if the linear forms $f_1, \dots, f_r$ are orthogonal in $E^\vee$.*

Through the choice of an $\mathcal{O}_K$-basis $e_1, \dots, e_n$ of $\Lambda$ we can identify $E$ with $K^n$. Then the dual lattice $\Lambda^\vee$ has the dual basis $e_1^\vee, \dots, e_n^\vee$ through which we can identify $E^\vee$ with $K^n$.

**Example 1.31.** Suppose $K = \mathbb{Q}_7$, $E = K^4$ and $\Lambda$ is the standard lattice $\mathbb{Z}_7^4$. The space $E$ is then a normed space with unit ball $\Lambda$. Set $A \in \mathbb{Q}_7^{4\times 4}$ to be the following matrix

$$A = \begin{bmatrix} 12 & 314 & 234 & 34 \\ 12 & 343 & 55 & 67 \\ 25 & 54 & 65 & 65 \\ 61 & 461 & 430 & 328 \end{bmatrix}.$$

The rows of $A$ define linear forms $f_1, \dots, f_4 \in E^\vee$ (the dual space $E^\vee$ is identified with $\mathbb{Q}_7^4$ thought the dual basis). Let us use Proposition 1.20 to test the orthogonality of the rows of $A$. Every row in this matrix has norm 1 in $E$ and modulo 7 the matrix becomes

$$\overline{A} = \begin{bmatrix} 5 & 6 & 3 & 6 \\ 5 & 0 & 6 & 4 \\ 4 & 5 & 2 & 2 \\ 5 & 6 & 3 & 6 \end{bmatrix} \in \mathbb{F}_7^{4\times 4}.$$

So if $Z = (Z_1, Z_2, Z_3, Z_4)^\top$ is a Gaussian vector with uniform distribution on $\mathbb{Z}_7^4$ and $Y = AZ$ we can see by Theorem 1.30 that $Y_1, Y_2, Y_3$ are independent but $Y_1, Y_2, Y_3, Y_4$ are not since the first three rows of $\overline{A}$ are linearly independent over $\mathbb{F}_7$ but the matrix $\overline{A}$ is singular.

## 1.2 A step into the Bruhat-Tits building

Buildings are geometric and combinatorial structures that generalize certain aspects of Riemannian symmetric spaces, see [94, 141]. These structures were introduced to understand reductive algebraic groups via their action which is of interest in geometric group theory [137, 136, 52, 55]. In this section we give a very brief introduction to the theory of buildings through the lattice class model for the Bruhat-Tits building of the reductive group PGL which will used repeatedly throughout this thesis. We refer the reader to [1] for more details on buildings.

Let $K$ be a discretely valuated field with surjective valuation $\mathrm{val} : K \to \mathbb{Z} \cup \{+\infty\}$. The *valuation ring* $\mathcal{O}_K$ is the set of elements in $K$ with non-negative valuation. Fix a uniformizer $\varpi$ of $K$. Then we denote by $k$ the residue field $k \coloneqq \mathcal{O}_K/\varpi\mathcal{O}_K$.

We fix an integer $d$ and consider the column space $V = K^d$. Since the ring $\mathcal{O}_K$ is a principal ideal domain, for any lattice $L$ in $V$, there exists a basis $(\epsilon_1, \ldots, \epsilon_d)$ of $V$ such that $L = \bigoplus_{i=1}^{d} \mathcal{O}_K \epsilon_i$. An example of a lattice is the *standard lattice* $L_0 = \mathcal{O}_K^d$ spanned by the standard basis $(e_1, \ldots, e_d)$. Two lattices $L_1$ and $L_2$ are called *equivalent* if there is some $u \in K^\times$ such that $L_2 = uL_1$. The *equivalence class* of $L$ (also referred to as the homothety class of $L$) is denoted by

$$[L] = \{uL : u \in K^\times\} = \{\varpi^m L : m \in \mathbb{Z}\}.$$

**Definition 1.32.** The *affine building* $\mathcal{B}_d(K)$ associated to the reductive algebraic group $\mathrm{PGL}(d, K)$ is a simplicial complex whose vertices are the equivalence classes $[L]$ of lattices $L$ in $E$. A set of vertices $\{[L_1], \ldots, [L_{s+1}]\}$ is an $s$-simplex in $\mathcal{B}_d(K)$ if and only if there exist representatives $\widetilde{L}_i \in [L_i]$ and a permutation $\sigma$ of $\{1, \ldots, s+1\}$ such that

$$\varpi\widetilde{L}_{\sigma(1)} \subsetneq \widetilde{L}_{\sigma(s+1)} \subsetneq \cdots \subsetneq \widetilde{L}_{\sigma(1)} \subsetneq \widetilde{L}_{\sigma(1)}.$$

The maximal simplices $\{[L_1], \ldots, [L_d]\}$ in $\mathcal{B}_d(K)$ are called *chambers*. Given a basis $g = (g_1, \ldots, g_d)$ of $V$, the *apartment* defined by $g \in \mathrm{GL}(d, K)$ is the set of homothety classes of lattices $[L]$ where $L$ is of the form

$$L = \varpi^{u_1}\mathcal{O}_K g_1 \oplus \cdots \oplus \varpi^{u_d}\mathcal{O}_K g_d, \quad \text{with } u_1, \ldots, u_d \in \mathbb{Z}.$$

The *standard apartment* is the apartment associated to the standard basis $(e_1, \ldots, e_d)$ of $V = K^d$. If $L$ is a lattice such that $[L]$ belongs to the standard apartment i.e.

$$L = \varpi^{u_1}\mathcal{O}_K e_1 \oplus \cdots \oplus \varpi^{u_d}\mathcal{O}_K e_d, \quad \text{for some } u_1, \ldots, u_d \in \mathbb{Z},$$

we call $u = (u_1, \ldots, u_d)$ the *exponent vector* of $L$ and write $L(u)$ to denote $L$.

**Remark 1.33.** Notice that if $\varpi L_1 \subsetneq L_s \subsetneq \cdots \subset L_1 \subsetneq L_1$ for lattices $L_1, \ldots, L_s$ then we have the following sequence of nested vector spaces over the residue field $k$

$$0 \subsetneq L_s/\varpi L_1 \subsetneq \cdots \subset L_1/\varpi L_1 \subsetneq L_1/\varpi L_1 \cong k^d.$$

Since $\dim_k(L_1/\varpi L_1) = d$, such a sequence has length as most $d$ so the maximal simplices in $\mathcal{B}_d(K)$ are of dimension $d-1$.

**Remark 1.34.** When $K$ is a local field, from Section 1.1.4, points of $\mathcal{B}_d(K)$ parametrize (up-to scaling) non-degenerate Gaussian measures on $V = K^d$. Points in the standard apartment parameterize (up-to scaling) Gaussian distributions $\mathbb{P}$ such that, if $X$ is a Gaussian vector distributed under $\mathbb{P}$, its coordinates in the standard basis are independent random variables.

The general linear group $\mathrm{GL}(d, K)$ acts (from the left) on $\mathcal{B}_d(K)$ preserving the simplicial structure [6]. This action is transitive on the lattice classes and also transitive on the chambers. The stabilizer of the lattice $L_0$ is the subgroup

$$\mathrm{GL}(d, \mathcal{O}_K) = \{g \in \mathcal{O}_K^{d\times d} : \mathrm{val}(\det(g)) = 0\}$$

and the stabilizer of the standard chamber is the *Iwahori* subgroup

$$B := \{g \in \mathrm{GL}(d, \mathcal{O}_K) : \mathrm{val}(g_{ij}) > 0 \text{ if } i < j\}.$$

Using the $\mathrm{GL}(d, K)$-action allows us to define the *type* of a lattice class. Fixing the standard lattice $L_0$ the type of $[gL_0]$ is $\mathrm{val}(\det(g)) + d\mathbb{Z} \in \mathbb{Z}/d\mathbb{Z} = \{0, 1, 2, \ldots, d-1\}$. In particular in the standard chamber, the type of $[L_i]$ is $i$ and in the standard apartment, the type of $L(u)$ is $u_1 + \ldots + u_d + d\mathbb{Z} \in \mathbb{Z}/d\mathbb{Z}$.

The type preserving automorphisms of $\mathcal{B}_d(K)$ form the group

$$\mathrm{VSL}_d(K) := \{g \in \mathrm{GL}_d(K) : \mathrm{val}(\det(g)) = 0\}.$$

Two chambers are called *adjacent*, if they share a common $d-2$-dimensional simplex.

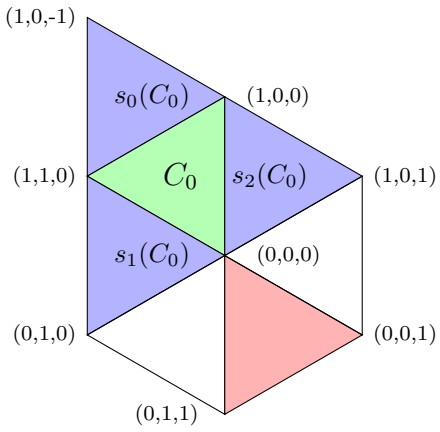



Figure 1.1: The standard chamber and its neighbors in the standard apartment in $\mathcal{B}_3(K)$.

The standard chamber is the simplex given with vertices $[M_0], \ldots, [M_{d-1}]$ defined by

$$M_i = \varpi\mathcal{O}_K e_1 \oplus \ldots \oplus \varpi\mathcal{O}_K e_i \oplus \mathcal{O}_K e_{i+1} \oplus \ldots \oplus \mathcal{O}_K e_d \quad \text{for } 1 \le i \le d-1$$

and

$$M_0 = \mathcal{O}_K e_1 \oplus \ldots \oplus \mathcal{O}_K e_d.$$

[6] Since the 0-simplices are homothety classes of lattices, this action factors through $\mathrm{PGL}(d, K)$.

Starting from the standard chamber $C_0$, there exist standard reflections $s_0, s_1, \ldots, s_{d-1}$ mapping $C_0$ to all the $d$ adjacent chambers in the standard apartment which are given as follows. For $i = 1, \ldots, d-1$ we define $s_i$ on the standard basis by

$$s_i(e_j) = \begin{cases} e_j & \text{if } j \neq i, i+1 \\ e_i & \text{if } j = i+1 \\ e_{i+1} & \text{if } j = i. \end{cases}$$

The map $s_0$ is defined by $s_0(e_i) = e_i$ for $i = 2, \ldots, d-1$ and $s_0(e_d) = \varpi e_1$, $s_0(e_1) = \varpi^{-1} e_d$. Then the set $\{s_0, \ldots, s_{d-1}\}$ forms a set of Coxeter generators for the affine Weyl group

$$W = \langle s_0, \ldots, s_{d-1} \rangle \leq \mathrm{VSL}_d(K).$$

This group $W$ acts regularly on the chambers in the standard apartment [22, § 1.5, Thm. 2], i.e. for every chamber $C$ in the standard apartment, there is a unique $w \in W$ such that $C = wC_0$. The group $W$ contains the group $D$ of all diagonal matrices $\mathrm{diag}(\varpi^{a_1}, \ldots, \varpi^{a_d})$ with $a_i \in \mathbb{Z}$ and $\sum_{i=1}^{d} a_i = 0$, as a normal subgroup. In fact, $W$ is isomorphic to the semidirect product of $D$ with the symmetric group $S_d$ of degree $d$; in symbols $W \cong D \rtimes S_d$. The famous Gauss-Bruhat decomposition then translates to

$$\mathrm{VSL}_d(K) = \bigcup_{w \in W} BwB.$$

We end this section by defining a notion of min and max convexity on the Bruhat-Tits building $\mathcal{B}_n(K)$. This notion was originally introduced by Faltings to study toroidal resolutions [70] and it appears in the context of Mustafin varieties [27, 85] (see also [169, Section 2.1]) and is tightly linked to tropical convexity in the sense of [96].

**Definition 1.35.** A set of vertices $S = \{[L_1], \ldots, [L_s]\}$ is *min-convex* (resp. *max-convex*) if for every pair of indices $1 \leq i \neq j \leq s$ and representatives $\widetilde{L}_i \in [L_i]$ and $\widetilde{L}_j \in [L_j]$, we have $[\widetilde{L}_i \cap \widetilde{L}_j] \in S$ ( resp. $[\widetilde{L}_i + \widetilde{L}_j] \in S$). The set $S$ is called *convex* if its is both min and max convex.

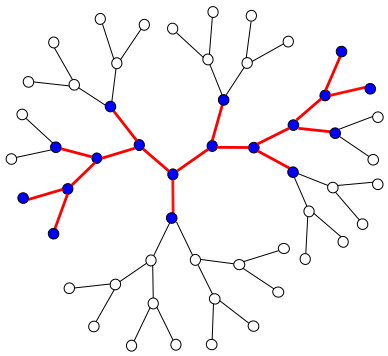


Figure 1.2: A convex set in the building $\mathcal{B}_2(\mathbb{Q}_2)$ (the set of vertices colored in blue).

## 1.3 Overview and contributions of this dissertation

Having introduced some necessary background, let us now give an overview of the main results of this dissertation. Our presentation is divided into two main parts: ***Probability*** and ***Number theory, combinatorics and geometry***.

### 1.3.1 Probability

The first part of this dissertation consists of chapters 2 to 4, which tackle problems of probabilistic nature. In more detail, this part consists of

1. Chapter 2 which is based on [54] is joint work with Ngoc Tran and on the sequel [48] to appear in the journal *Algebraic statistics*. In this chapter, we study the entropy map for multivariate Gaussian distributions on a non-archimedean local field. As in the real case, the image of this map lies in the supermodular cone. Moreover, given a multivariate Gaussian measure on a local field, its image under the entropy map determines its pushforward under the valuation map. In general, this map can be defined for non-archimedian valued fields whose valuation group is an additive subgroup of the real line, and it remains supermodular. We also explicitly compute the image of this map in dimension 3 and discuss a several non-archimedean statistical Gaussian models.

2. Chapter 3 which is based on joint work [50] with Enis Kaya. In this chapter, we present a method to sample from algebraic manifolds (embedded in some projective or affine space) defined over a local field. Our method is inspired but the work of Breiding and Marigliano [23] on sampling from real algebraic manifolds. In a nutshell, this method is based on slicing the manifold in question with random linear spaces of complementary dimension, then sampling from the resulting finite intersection using a suitable weight function.

3. Chapter 4 which is based on [53] which is joint work with Jim Pitman. In this chapter, we introduce a novel characterization of Bernoulli polynomials by circular convolution. There is a combinatorial and probabilistic model underlying this characterization which we call *The Bernoulli clock*. We use this model to give a probabilistic perspective to the work of Horton and Kurn [91] and the more recent work of Clifton et al. [33] on counting permutations of the multiset $1^m \dots n^m$ with longest continuous and increasing subsequence of length $n$ starting from 1.

### 1.3.2 Number theory, combinatorics and geometry

The second part of this dissertation consists of chapters 5 to 7 which tackle problems in number theory, combinatorics and geometry. In more detail, this part consists of

1. Chapter 5 which is based on joint work [55] with Marvin A. Hahn, Gabriele Nebe, Mima Stanojkovski and Bernd Sturmfels, and its sequel [52], which is joint work with Gabriele Nebe and Mima Stanojkovski. In this chapter, we study orders in the ring of matrices over a non-archimedean discretely valued field and their action of the Bruhat-Tits building. Namely, we study *graduated orders* in the sense of [133] using the modern language of tropical geometry, and extend our study to a bigger class of orders for which we coined the term *bolytrope orders*. In particular, we describe the set of fixed points of these orders in the Bruhat-Tits building.

2. Chapter 6 which is based joint work [51] with Antonio Lerario. In this chapter, given a non-archimedean discretely valued field $K$ whose valuation ring is $R$, an vector space $V$ of dimension $n$ over $K$ and a partition $\lambda$ of a positive integer $d$, we study the problem of determining the set of $\mathrm{GL}(n, R)$-invariant lattices in the Schur module $S_\lambda(V)$. This chapter leverages some results that are established in Chapter 5.

3. Chapter 7 which is based in joint work [49] with Paul A. Helminck and Enis Kaya. In this chapter we express the reduction types of Picard curves in terms of tropical invariants associated to binary quintics. These invariants are connected to Picard modular forms using recent work by Cléry and van der Geer. We furthermore give a general framework for tropical invariants associated to group actions on arbitrary varieties. The problem of describing reduction types of curves in terms of their associated invariants fits in this general framework by mapping the space of binary forms to symmetrized versions of the Deligne–Mumford compactification $\bar{M}_{0,n}$. We conjecture that the techniques introduced here can be used to find tropical invariants for binary forms of any degree.

# Part I

# Probability

# Chapter 2

# Entropy and statistics of Gaussian measures on local fields.

This chapter is based on joint work [54] with Ngoc Tran and the sequel [48]. This chapter is to appear in modified form in the journal *Algebraic statistics*.

## 2.1 Introduction and notation

Gaussian measures on local fields are introduced in [65]. In this text, we aim to exhibit the entropy map of these measures and discuss the properties this map satisfies. Our aim is to highlight the similarities with the real case. Before we discuss Gaussian measures on local fields (see Section 2.2), we begin by reviewing the entropy map in the real setting.

### 2.1.1 Entropy of real multivariate Gaussian distributions

For a positive integer $d$, multivariate Gaussian distributions on $\mathbb{R}^d$ are determined by their mean $\mu \in \mathbb{R}^d$ and their positive semi-definite covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$. Hence the natural parameter space for *centered* (i.e with zero mean) Gaussian distributions on $\mathbb{R}^d$ is the positive semi-definite cone in $\mathbb{R}^{d\times d}$, which we denote by

$$\mathrm{PSD}_d := \{\Sigma \in \mathrm{Sym}_d(\mathbb{R})\colon \langle x, \Sigma x\rangle \geq 0 \text{ for all } x \in \mathbb{R}^d\},$$

where $\mathrm{Sym}_d(\mathbb{R})$ is the space of real symmetric matrices in $\mathbb{R}^{d\times d}$ and $\langle\cdot,\cdot\rangle$ is the usual inner product on $\mathbb{R}^d$. *Non-degenerate Gaussian* distributions are those whose covariance matrix $\Sigma$ is positive definite, i.e, $\Sigma \in \mathrm{PD}_d$ where

$$\mathrm{PD}_d := \mathrm{PSD}_d^\circ = \{\Sigma \in \mathrm{Sym}_d(\mathbb{R})\colon \langle x, \Sigma x\rangle > 0 \text{ for all non zero } x \in \mathbb{R}^d\}.$$

There is no shortage of instances where the PSD cone appears in probability and statistics [156], optimization [122, Chapter 12] and combinatorics [79].

The positive definite cone has a pleasant group-theoretic structure in the sense that its elements are in one-to-one correspondence with left cosets of the orthogonal group $\mathrm{O}(d,\mathbb{R})$ in the general linear group $\mathrm{GL}(d,\mathbb{R})$. The map sending the coset $A\mathrm{O}(d,\mathbb{R}) \in \mathrm{GL}(d,\mathbb{R})/\mathrm{O}(d,\mathbb{R})$ to $AA^T \in \mathrm{PD}_d$ is a bijection. This underscores the fact that multivariate Gaussians are tightly linked to the linearity and orthogonality structures that the Euclidean space $\mathbb{R}^d$ enjoys.

An important concept in statistics, probability, and information theory is the notion of entropy, which is a measure of uncertainty and disorder in a distribution; see [120]. The entropy of a centered multivariate Gaussian with covariance matrix $\Sigma$ is given, up to an additive constant, by

$$h(\Sigma) = -\log(|\det(\Sigma)|) = -\log(\det\Sigma).$$

If $X$ is a random vector in $\mathbb{R}^d$ with non-degenerate centered Gaussian distribution given by a covariance matrix $\Sigma \in \mathrm{PD}_d$, then for any subset $I$ of $[d] \coloneqq \{1,2,\ldots,d\}$ the vector $X_I$ of coordinates of $X$ indexed by $I$ is also a random vector with non-degenerate Gaussian measure on $\mathbb{R}^{|I|}$. Moreover, its covariance matrix is $\Sigma_I = (\Sigma_{i,j})_{i,j\in I} \in \mathbb{R}^{|I|\times|I|}$, so we can define the entropy $h_I(\Sigma)$ of $X_I$ as

$$h_I(\Sigma) \coloneqq h(\Sigma_I) = -\log(\det(\Sigma_I)).$$

The collection of entropy values $(h_I(\Sigma))_{I\subset[d]}$ satisfies the inequalities

$$h_I(\Sigma) + h_J(\Sigma) \le h_{I\cap J}(\Sigma) + h_{I\cup J}(\Sigma) \text{ for any two subsets } I,J\subset[d]. \tag{2.1}$$

This is thanks to what is known as Koteljanskii's inequalities [103] on the determinants of positive definite matrices, i.e,

$$\det(\Sigma_I)\ \det(\Sigma_J) \quad\geq\quad \det(\Sigma_{I\cap J})\ \det(\Sigma_{I\cup J}). \tag{2.2}$$

In the language of polyhedral geometry this means that the image of the entropy map

$$\begin{aligned} H : \mathrm{PD}_d &\to \mathbb{R}^{2^d} \\ \Sigma &\mapsto (h_I(\Sigma))_{I\subset[d]} \end{aligned} \tag{2.3}$$

lies inside the *supermodular* cone $\mathcal{S}_d$ in $\mathbb{R}^{2^d}$. This is the *polyhedral cone* specified by the inequalities in (2.1), i.e,

$$\mathcal{S}_d \coloneqq \{x = (x_I)_{I\subset[d]} \in \mathbb{R}^{2^d} : x_\emptyset = 0 \text{ and } x_I + x_J \le x_{I\cap J} + x_{I\cup J} \text{ for all } I,J\subset[d]\}.$$

Since $x_\emptyset = 0$ for $x \in \mathcal{S}_d$ we can see $\mathcal{S}_d$ as a cone in $\mathbb{R}^{2^d-1}$.

### 2.1.2 Main results

In this chapter we deal with multivariate Gaussian distributions on local fields as defined in Section 1.1.4, and more generally nonarchimedean valued fields. See Example 2.3 for a discussion. In particular we shall define an analog to the entropy map and show that it satisfies the same set of inequalities (2.1). More precisely we prove the following:

**Theorem 2.1.** *The push-forward measure of a multivariate Gaussian measure on a local field by the valuation map is given by a tropical polynomial whose coefficients are determined by the entropy map of this measure (see Theorem 2.9). Moreover, these coefficients are supermodular. The entropy map is still well-defined on non-archimedian valued fields in general, and remains supermodular (see Theorem 2.13).*

One motivation behind this chapter is the search for a suitable definition of *tropical Gaussian measures* [162]. Tropical stochastics has been an active research area in the recent years and has diverse applications from phylogenetics [113, 167] to game theory [6] and economics [15, 163]. One appealing approach to define *tropical Gaussians* is to tropicalize Gaussian measures on a valued field.

Our text is organized as follows. In Section 2.3 we show that tropicalizing multivariate Gaussians on local fields yields probability measures on the lattice $\mathbb{Z}^d$ that are determined by the entropy map via a tropical polynomial. In Section 2.3.1 we show the supermodularity of the entropy map and provide a recursive algorithm to compute it. In Section 2.4, we explain why orthogonality is not a suitable approach to define Gaussian measures when the field $K$ is not locally compact. Nevertheless, we will see that the entropy map is still well defined and remains supermodular and we explicitly compute its image when $d = 3$.

Implementations, computations and data related to this chapter are made available at

`https://mathrepo.mis.mpg.de/GaussianEntropyMap/index.html`. (2.4)

**Remark 2.2.** For readers not familiar with local fields, we refer to Section 1.1.2 or to [100, 146] for a more detailed treatment. Local fields are not commonly used in statistics and probability. However, in recent years there has been a stream of literature addressing probabilistic and statistical questions in the $p$-adic setting, starting from the early work of Evans [60, 65] to the more recent developments [28, 67, 105] to mention a few.

## 2.2 Background on valued fields and Gaussian measures

This section is meant to establish some notation and recall the basic facts and result that we will need in our discussion. Most of these results can be found in the literature on valued fields in number theory [57, 146, 166] and functional analysis [140, 144] (see also Chapter 1 for a quick introduction).

### 2.2.1 Valued fields

Let $K$ be a field with an *additive nonarchimedean valuation* $\mathrm{val} : K \to \mathbb{R} \cup \{+\infty\}$ with valuation group $\Gamma := \mathrm{val}(K^\times)$. The valuation map val defines an equivalence class of *exponential valuations* or *absolute values* $|\cdot|$ on $K$ via $|x| := a^{-\mathrm{val}(x)}$ (where $a \in (1, \infty)$) and hence also a topology on $K$. The valuation val is called *discrete* if its valuation group $\Gamma$ is a discrete subgroup of $\mathbb{R}$ which, by scaling val suitably, we can always assume that $\Gamma = \mathbb{Z}$ (we then call val a *normalized valuation*). In the discrete valuation case we fix a uniformizer $\varpi$ of $K$, i.e, an element $\varpi \in K$ with $\mathrm{val}(\varpi) = 1$. We denote by $\mathcal{O} := \{x \in K, \mathrm{val}(x) \geq 0\}$ the valuation ring of $K$; this is a local ring with unique maximal ideal $\mathfrak{m} := \{x \in K, \mathrm{val}(x) > 0\}$ and residue field $k := \mathcal{O}/\mathfrak{m}$. When the valuation is discrete, the ideal $\mathfrak{m}$ is generated in $\mathcal{O}$ by $\varpi$ i.e $\mathfrak{m} := \varpi\mathcal{O}$. We mention typical examples of such fields in Example 2.3.

**Example 2.3.** (1) The field $\mathbb{F}_q((t))$ of Laurent series in one variable with coefficients in the finite field $\mathbb{F}_q$.

(2) The fields $\mathbb{R}((t))$ or $\mathbb{C}((t))$ of Laurent series with complex or real coefficients. These are fields with an infinite residue field but still in discrete valuation $\Gamma = \mathbb{Z}$.

(3) The fields $\mathbb{R}\{\{t\}\} = \cup_{n\geq 1}\mathbb{R}((t^{1/n}))$ and $\mathbb{C}\{\{t\}\} = \cup_{n\geq 1}\mathbb{C}((t^{1/n}))$ of Puiseux series in $t$. In this case the valuation group $\Gamma = \mathbb{Q}$ is dense.

(4) The field of generalized Puiseux series $\mathbb{K}$ which has valuation group $\Gamma = \mathbb{R}$. This field consists of formal series $\boldsymbol{f} = \sum_{\alpha\in\mathbb{R}} a_\alpha t^\alpha$ where $\mathrm{supp}(\boldsymbol{f}) := \{\alpha \in \mathbb{R} : a_\alpha \neq 0\}$ is either finite or has $+\infty$ as the only accumulation point; see [7].

(5) All the previous fields have the same characteristic as their residue fields. Interesting examples in mixed characteristic are the field of $p$-adic numbers $\mathbb{Q}_p$ where $p$ is prime, its algebraic closure $\overline{\mathbb{Q}}_p$ and the field of $p$-adic complex numbers $\mathbb{C}_p$ (completion of $\overline{\mathbb{Q}}_p$).

#### 2.2.1.1 Local fields

These are valued fields that are locally compact. In this section let us assume that $K$ is locally compact. It is then known that $K$ is isomorphic to a finite field extension of $\mathbb{Q}_p$ or $\mathbb{F}_q((t))$ and that its valuation group $\Gamma$ is discrete in $\mathbb{R}$, and its residue field $k$ is finite. In this case, by convention, the absolute value on $K$ is defined as $|x| = q^{-\mathrm{val}(x)}$ (so we choose $a = q$), and there exists a unique Haar measure $\mu$ on $K$ such that $\mu(\mathcal{O}) = 1$.

### 2.2.2 Lattices

Let $d \geq 1$ an integer. We call a *lattice* in $K^d$ any $\mathcal{O}$-submodule $\Lambda := \bigoplus_{i=1}^d \mathcal{O}a_i$ generated by a basis $(a_1, \ldots, a_d)$ of $K^d$. The basis $(a_1, \ldots, a_d)$ that generates $\Lambda$ is not unique. We can write $\Lambda = A\mathcal{O}^d$ where $A$ is the matrix with columns $a_1, \ldots, a_d$, which is then called a *representative* of $\Lambda$. The elements $U$ of the group $\mathrm{GL}(d, K)$ that leave $\mathcal{O}^d$ invariant (i.e $U\mathcal{O}^d = \mathcal{O}^d$) are exactly the matrices $U \in \mathrm{GL}(d, \mathcal{O})$ with entries in $\mathcal{O}$ whose inverse has all

entries in $\mathcal{O}$. The group $\mathrm{GL}(d,\mathcal{O})$ then plays the role of the orthogonal group $\mathrm{O}(d,\mathbb{R})$ [69, Theorem 2.4]. Then, like positive definite matrices, lattices are in one-to-one correspondence with left cosets $\mathrm{GL}(d,K)/\mathrm{GL}(d,\mathcal{O})$, in particular, any two representatives of a lattice $\Lambda$ are elements of the same left coset. A lattice $\Lambda$ is called *diagonal*[1] if it admits a diagonal matrix as a representative. Let us now state a result on lattices over valued fields that will be useful in our discussion.

**Lemma 2.4.** *For any two lattices $\Lambda, \Lambda'$ there exists an element $g \in \mathrm{GL}(d,K)$ such that $g\Lambda$ and $g\Lambda'$ are both diagonal lattices.* [2]

*Proof.* It suffices to show this when $\Lambda$ is the standard lattice $\Lambda = \mathcal{O}^d$. Let $A \in \mathrm{GL}(d,K)$ be a representative of $\Lambda'$. Thanks to the nonarchimedean singular value decomposition (see Proposition 1.24), there exists a diagonal matrix $D \in \mathrm{GL}(d,K)$ and $U, V \in \mathrm{GL}(d,\mathcal{O})$ such that $A = UDV$. Hence we deduce that $\Lambda' = UD\mathcal{O}^d$. Picking $g = U^{-1}$ yields $g\Lambda = U^{-1}\mathcal{O}^d = \mathcal{O}^d$ and $g\Lambda' = D\mathcal{O}^d$. □

#### 2.2.2.1 Gaussian measures

Suppose that $K$ is a local field and $d$ is a positive integer. As discussed in Section 2.2, one can define multivariate *Gaussian* measures on $K^d$ using nonarchimedean orthogonality. It turns out that these measures are precisely the uniform distributions on $\mathcal{O}$-submodules of $K^d$. The non-degenerate Gaussians on $K^d$ are then parameterized by full rank submodules of $K^d$ i.e. lattices.

For a lattice $\Lambda$ in $K^d$ we denote by $\mathbb{P}_\Lambda$ the Gaussian measure on $K^d$ given by $\Lambda$, i.e. the uniform probability measure on $\Lambda$. If $f_\Lambda$ denotes the density (with respect to the Haar measure $\mu^{\otimes d}$) of $\mathbb{P}_\Lambda$, then

$$f_\Lambda(x) = \mathbf{1}_\Lambda(x)/\mu^{\otimes d}(\Lambda), \quad x \in K^d,$$

where $\mathbf{1}_\Lambda$ is the set indicator function of $\Lambda$.

One can then think of lattices as analogs for the positive definite covariance matrices in the real case since they parameterize non-degenerate multivariate Gaussian measures. In the language of group theorists, one can think of the Bruhat-Tits building for the reductive group $\mathrm{PGL}_d(K)$ [1] as the parameter space for non-degenerate Gaussians up to scalar multiplication.

## 2.3 The entropy map of local field Gaussian distributions

In this section we assume that $K$ is a local field and we fix a positive integer $d \geq 1$ and a lattice $\Lambda$ in $K^d$. We recall that there is a unique Haar measure $\mu^{\otimes d}$ on $K^d$ which is the

[1] homothety classes of diagonal lattices form what is called an *apartment* in the theory of buildings.
[2] This is in fact a property of buildings: any two chambers belong to a common apartment. See [1].

product measure induced by $\mu$ on $K$. Letting $A$ be a representative of the lattice $\Lambda$, i.e. $\Lambda = A\mathcal{O}^d$, we can define the *entropy* $h(\Lambda)$ of the lattice $\Lambda$ as

$$h(\Lambda) = \mathrm{val}(\det(A)).$$

This is a well defined quantity since any other representative of $\Lambda$ is of the form $AU$ where $U \in \mathrm{GL}(d, \mathcal{O})$ and $\det(U) \in \mathcal{O}^\times$ is a unit, so $\mathrm{val}(\det(U)) = 0$. This definition lines up with the definition in the real case because $\mathrm{val}(x) = -\log_q(|x|)$ where $|\cdot|$ is the absolute value on $K$, so we get

$$h(\Lambda) = \mathrm{val}(\det(A)) = -\log_q(|\det(A)|).$$

The following proposition justifies the nomenclature "entropy" and relates the entropy $h(\Lambda)$ of a lattice $\Lambda$ to its measure $\mu^{\otimes d}(\Lambda)$.

**Proposition 2.5.** *We have $\mu^{\otimes d}(\Lambda) = q^{-h(\Lambda)}$. Moreover, the quantity $h(\Lambda)$ is the differential entropy of the Gaussian measure $\mathbb{P}_\Lambda$, i.e,*

$$h(\Lambda) = \int_{K^d} \log_q(f_\Lambda(x))\mathbb{P}_\Lambda(dx).$$

*Proof.* Let $A$ be a representative of $\Lambda$. Thanks to the nonarchimedean singluar value decomposition (see [62, Theorem 3.1]), we can write $A = UDV$, where $U, V \in \mathrm{GL}(d, \mathcal{O})$ are two orthogonal matrices and $D$ is a diagonal matrix. Then we have $\Lambda = UD\mathcal{O}^d$. Since orthogonal linear transformations in $K^d$ preserve the measure, we have $\mu^{\otimes d}(\Lambda) = \mu^{\otimes}(D\mathcal{O}^d)$. Let $\alpha_1, \dots, \alpha_d$ be the diagonal entries of $D$. Then we have $\mu^{\otimes d}(\Lambda) = \mu^{\otimes d}(\bigoplus_{i=1}^d \alpha_i\mathcal{O}) = q^{-\mathrm{val}(\alpha_1)-\cdots-\mathrm{val}(\alpha_d)}$. But $\mathrm{val}(\alpha_1) + \cdots + \mathrm{val}(\alpha_d) = \mathrm{val}(\det(A)) = h(\Lambda)$. The second statement follows from the immediate computation:

$$\int_{K^d} \log_q(f_\Lambda(x))\mathbb{P}_\Lambda(dx) = \int_{K^d} \log_q(f_\Lambda(x))f_\Lambda(x)\mu^{\otimes d}(dx) = h(\Lambda). \qquad \square$$

For a subset $I$ of $[d] \coloneqq \{1, 2, \dots, d\}$ we denote by $\Lambda_I$ the image of $\Lambda$ under the projection onto the space $K^{|I|}$ of coordinates indexed by $I$. This is also a lattice in the space $K^{|I|}$. So, for any subset $I \subset [d]$, we can define the entropy $h_I(\Lambda)$ of the lattice $\Lambda_I$. We can then define the entropy map

$$H : \mathrm{GL}(d, K)/\mathrm{GL}(\mathcal{O}) \to \mathbb{R}^{2^d}, \quad \Lambda \mapsto (h_I(\Lambda))_{I\subset[d]}, \tag{2.5}$$

where $h_\emptyset(\Sigma) = 0$ by convention. If $A$ is a representative of $\Lambda$ with columns $a_1, \dots, a_d$, then the lattice $\Lambda_I$ is the lattice generated over $\mathcal{O}$ by the vectors $a_{i,I}$ which are the sub-vectors of the $a_i$'s with coordinates indexed by $I$. So we can compute $h_I(\Lambda)$ from the matrix $A$ by

$$h_I(\Lambda) = \min_{J\subset[d],|J|=|I|} \mathrm{val}(\det(A_{I\times J})), \tag{2.6}$$

where $A_{I\times J}$ is the matrix extracted from $A$ by taking the rows indexed by $I$ and the columns indexed by $J$, i.e. $A_{I\times J} = (A_{i,j})_{i\in I, j\in J}$.

Now let $X$ be a $K^d$-valued random variable with Gaussian distribution $\mathbb{P}_\Lambda$ given by $\Lambda$. So for any measurable set $B$ in the Borel $\sigma$-algebra of $K^d$,

$$\mathbb{P}_\Lambda(X \in B) = \frac{\mu^{\otimes d}(\Lambda \cap B)}{\mu^{\otimes d}(\Lambda)},$$

and $V \coloneqq \mathrm{val}(X)$ its image under coordinate-wise valuation. Notice that, since $\mathbb{P}_\Lambda(X_i = 0) = 0$ for any $i \in \{1, \dots, d\}$, the vector $V$ is almost surely in $\mathbb{Z}^d$ . By definition, the distribution of $V$ is the push-forward of the distribution of $X$ by the map val. We are interested in the distribution of the valuation vector $V$ and to determine it we compute its *tail distribution function* $Q_\Lambda$ which is defined on $\mathbb{R}^d$ as

$$Q_\Lambda(v) \coloneqq \mathbb{P}_\Lambda(V \geq v) \text{ for any } v \in \mathbb{R}^d,$$

where $\geq$ is the coordinate-wise partial order on $\mathbb{R}^d$. Since $V$ takes values in $\mathbb{Z}^d$, this function is completely determined by its values for $v \in \mathbb{Z}^d$. For a vector $v = (v_1, \dots, v_d) \in \mathbb{Z}^d$ let us denote by $\boldsymbol{\varpi}^{\boldsymbol{v}}$ the $\mathcal{O}$-module generated by the basis $\varpi^{v_i} e_i$ where $e_1, \dots, e_d$ is the standard basis of $K^d$ i.e.

$$\boldsymbol{\varpi}^{\boldsymbol{v}} = \varpi^{v_1}\mathcal{O}e_1 \oplus \cdots \oplus \varpi^{v_d}\mathcal{O}e_d.$$

**Definition 2.6.** We define the *logarithmic tail distribution function* $\varphi_\Lambda$ as

$$\varphi_\Lambda \colon \mathbb{Z}^d \to \mathbb{Z}, \quad v \mapsto -\log_q(Q_\Lambda(v)).$$

The following lemma relates the tail distribution function $\varphi_\Lambda$ with the entropy $h(\Lambda)$ of the lattice $\Lambda$.

**Lemma 2.7.** *We have* $\varphi_\Lambda(v) = h(\Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\Lambda)$. *Moreover, if* $[\Lambda : \Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}]$ *denotes the index of* $\Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}$ *as a subgroup of* $\Lambda$ *then we also have*

$$Q_\Lambda(v) = 1/[\Lambda : \Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}].$$

*Proof.* By definition we have $Q_\Lambda(v) = \mathbb{P}_\Lambda(X \in \boldsymbol{\varpi}^{\boldsymbol{v}}) = \mu^{\otimes d}(\boldsymbol{\varpi}^{\boldsymbol{v}} \cap \Lambda)/\mu^{\otimes d}(\Lambda)$. So by virtue of Proposition 2.5 we deduce that $Q_\Lambda(v) = q^{h(\Lambda) - h(\Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}})}$. The first statement then follows from the definition of $\varphi_\Lambda$ (Definition 2.6). For the second statement, by definition, $\Lambda$ can be partitioned into $[\Lambda : \Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}]$ cosets of $\Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}$. Since the Haar measure $\mu^{\otimes d}$ is translation invariant all of these cosets have the same measure i.e. $\mu^{\otimes d}(\Lambda) = [\Lambda : \Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}]\mu^{\otimes d}(\Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}})$. The result then follows from the fact that $Q_\Lambda(v) = \mu^{\otimes d}(\boldsymbol{\varpi}^{\boldsymbol{v}} \cap \Lambda)/\mu^{\otimes d}(\Lambda)$ and Definition 2.6. □

Next, we introduce a technical tool that we will be using in the proof of our first result.

**Definition 2.8.** For any $\ell \in \{0, \dots, d\}$ we define the *$\ell$-distance* $\phi_\ell(\Lambda, \Lambda')$ of two lattices $\Lambda, \Lambda'$ as the minimum of $\mathrm{val}(\det(x_1, \dots, x_\ell, y_1, \dots, y_k))$ among all possible choices of $x_1, \dots, x_\ell \in \Lambda$ and $y_1, \dots, y_k \in \Lambda'$ where $k = d - \ell$.

Since for any $g \in \mathrm{GL}(d,K)$ , $x_1, \dots x_\ell \in \Lambda$ and $y_1, \dots, y_k \in \Lambda'$ we have

$$\mathrm{val}(\det(gx_1, \dots gx_\ell, gy_1, \dots, gy_k)) = \mathrm{val}(\det(x_1, \dots, x_\ell, y_1, \dots, y_k)) + \mathrm{val}(\det(g)),$$

we can see that

$$\phi_\ell(g.\Lambda, g.\Lambda') = \phi_\ell(\Lambda, \Lambda') + \mathrm{val}(\det(g)).$$

We then deduce that the quantity $\phi_\ell(\Lambda, \Lambda') - h(\Lambda')$ is invariant under the action $\mathrm{GL}(d,K)$; that is for any $g \in \mathrm{GL}(d,K)$ we have

$$\phi_\ell(g.\Lambda, g.\Lambda') - h(g.\Lambda') = \phi_\ell(\Lambda, \Lambda') - h(\Lambda').$$

When the second lattice $\Lambda' = \boldsymbol{\varpi}^{\boldsymbol{v}}$ is diagonal and $\Lambda$ has representative $A \in \mathrm{GL}(d,K)$, the optimal choice for the vectors $x_1, \dots, x_\ell$ and $y_1, \dots, y_k$ is when the vectors $x_1, \dots, x_\ell$ are among the columns $a_1, \dots, a_d$ of $A$ and the vectors $y_1, \dots, y_k$ are among the vectors $\varpi^{v_i} e_i$ where $(e_i)_{1 \le i \le d}$ is the standard basis of $K^d$. So we deduce that $\phi_\ell(\Lambda, \boldsymbol{\varpi}^{\boldsymbol{v}})$ can be computed as follows:

$$\phi_\ell(\Lambda, \boldsymbol{\varpi}^{\boldsymbol{v}}) = \min_{\substack{I,J \subset [d] \\ |I|=|J|=\ell}} \left( \mathrm{val}(\det(A_{I \times J})) + \sum_{j \notin J} v_j \right).$$

So we also get

$$\phi_\ell(\Lambda, \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\boldsymbol{\varpi}^{\boldsymbol{v}}) = \min_{\substack{I,J \subset [d] \\ |I|=|J|=\ell}} \left( \mathrm{val}(\det(A_{I \times J})) - \sum_{j \in J} v_j \right). \tag{2.7}$$

In the special case $\Lambda = \boldsymbol{\varpi}^{\boldsymbol{a}}$, for $a \in \mathbb{Z}^d$, the determinant of $A_{I \times J}$ in the above optimization problem is 0 whenever $J \neq I$, since we can choose $A$ to be diagonal. So we get

$$\phi_\ell(\boldsymbol{\varpi}^{\boldsymbol{a}}, \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\boldsymbol{\varpi}^{\boldsymbol{v}}) = \min_{I \subset [d], |I|=\ell} \left( \sum_{i \in I} a_i - \sum_{i \in I} v_i \right).$$

**Theorem 2.9.** *The logarithmic tail distribution function $\varphi_\Lambda$ is the tropical polynomial on $\mathbb{Z}^d$ given by*

$$\varphi_\Lambda(v) = \max_{I \subset [d]} (v_I - h_I(\Lambda)). \tag{2.8}$$

*Proof.* First we show this for a diagonal lattice $\Lambda = \boldsymbol{\varpi}^{\boldsymbol{a}}$ where $a \in \mathbb{Z}^d$. For any $v \in \mathbb{Z}^d$, let $a \vee v$ the vector with coordinates $\max(a_i, v_i)$. We have $\boldsymbol{\varpi}^{\boldsymbol{a}} \cap \boldsymbol{\varpi}^{\boldsymbol{v}} = \boldsymbol{\varpi}^{\boldsymbol{a} \vee \boldsymbol{v}}$ so we get the entropy $h(\boldsymbol{\varpi}^{\boldsymbol{a}}) = \sum_{i=1}^d a_i$ and $h(\boldsymbol{\varpi}^{\boldsymbol{a}} \cap \boldsymbol{\varpi}^{\boldsymbol{v}}) = h(\boldsymbol{\varpi}^{\boldsymbol{a} \vee \boldsymbol{v}}) = \sum_{i=1}^d \max(a_i, v_i)$. Hence we have

$$\varphi_\Lambda(v) = h(\boldsymbol{\varpi}^{\boldsymbol{a}} \cap \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\boldsymbol{\varpi}^{\boldsymbol{a}}) = \max_{I \subset [d]} \left( \sum_{i \in I} v_i + \sum_{i \notin I} a_i \right) - \sum_{i=1}^d a_i = \max_{I \subset [d]} (v_I - a_I),$$

and $h_I(\boldsymbol{\varpi}^{\boldsymbol{a}}) = a_I$. So the theorem holds for diagonal lattices. To see why it also holds for a general lattice $\Lambda$, first notice that in the diagonal case $\Lambda = \boldsymbol{\varpi}^{\boldsymbol{a}}$ we have

$$\varphi_\Lambda(v) = -\min_{\ell=0,\dots,d} \left(\phi_\ell(\Lambda, \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\boldsymbol{\varpi}^{\boldsymbol{v}})\right).$$

Secondly, notice that the right hand side of the previous equation is invariant under the action of $\mathrm{GL}(d, K)$. So for $g \in \mathrm{GL}(d, K)$,

$$\min_{\ell=0,\dots,d} \left(\phi_\ell(g.\Lambda, g.\boldsymbol{\varpi}^{\boldsymbol{v}}) - h(g.\boldsymbol{\varpi}^{\boldsymbol{v}})\right) = \min_{\ell=0,\dots,d} \left(\phi_\ell(\Lambda, \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\boldsymbol{\varpi}^{\boldsymbol{v}})\right).$$

By Lemma 2.7, we have $\varphi_\Lambda(v) = \log_q([\Lambda : \Lambda \cap \boldsymbol{\varpi}^{\boldsymbol{v}}]) = \log_q([g.\Lambda : g.\Lambda \cap g.\boldsymbol{\varpi}^{\boldsymbol{v}}])$. Now fix a general lattice $\Lambda$ and $v \in \mathbb{Z}^d$. Also, by Lemma 2.4, there exists $g \in \mathrm{GL}(d, K)$ such that $g\Lambda$ and $g\boldsymbol{\varpi}^{\boldsymbol{v}}$ are both diagonal, so

$$\begin{aligned}\varphi_\Lambda(v) = \log_q([g.\Lambda : g.\Lambda \cap g.\boldsymbol{\varpi}^{\boldsymbol{v}}]) &= -\min_{\ell=0,\dots,d} \left(\phi_\ell(g.\Lambda, g.\boldsymbol{\varpi}^{\boldsymbol{v}}) - h(g.\boldsymbol{\varpi}^{\boldsymbol{v}})\right) \\ &= -\min_{\ell=0,\dots,d} \left(\phi_\ell(\Lambda, \boldsymbol{\varpi}^{\boldsymbol{v}}) - h(\boldsymbol{\varpi}^{\boldsymbol{v}})\right).\end{aligned}$$

Hence, we deduce, thanks to equation (2.7), that

$$\varphi_\Lambda(v) = -\min_{\ell=0,\dots,d} \left( \min_{\substack{I,J\subset[d] \\ |I|=|J|=\ell}} \left( \mathrm{val}(\det(A_{I\times J})) - \sum_{j\in J} v_j \right) \right).$$

We can simplify this thanks to equation (2.6) to get the desired equation (2.8). □

So the distribution of the random vector of valuations $V$ is given by a tropical polynomial $\varphi_\Lambda$ via its tail distribution function $Q_\Lambda$. The coefficients of this polynomial are exactly the entropies $h_I(\Lambda)$. Now we prove a couple of interesting properties of $\varphi_\Lambda$, namely how the coefficients $h_I(\Lambda)$ behave under diagonal scaling and permutation of coordinates of the random vector $X$. To this end, let us denote by $D_a = \mathrm{diag}(a_1, \dots, a_n)$ the diagonal matrix with coefficients $a_i \in K$ and $P^\sigma$ the permutation matrix corresponding to a permutation $\sigma$ of $[d]$ i.e $P^\sigma_{i,j} = 1$ when $j = \sigma(i)$ and 0 otherwise.

**Lemma 2.10.** *Let $\Lambda$ be a lattice in $K^d$, $a \in K^d$ and $\sigma$ a permutation of $[d]$. We have*

$$h_I(D_a\Lambda) = h_I(\Lambda) + \sum_{i\in I} \mathrm{val}(a_i) \text{ and } h_I(P^\sigma\Lambda) = h_{\sigma(I)}(\Lambda).$$

*Proof.* For $I \subset [d]$, we have $h_I(D_a\Lambda) = \min_{|J|=|I|} \mathrm{val}(\det((D_aA)_{I\times J}))$, where $A$ is any representative of $\Lambda$. Since all the rows of $D_aA$ are multiples of those of $A$ by the scalars $a_i$ we deduce that $\det((D_aA)_{I\times J}) = \det(A_{I\times J}) \prod_{i\in I} a_i$ and hence we get

$$h_I(D_a\Lambda) = h_I(\Lambda) + \sum_{i\in I} \mathrm{val}(a_i).$$

Similarly we can see the effect the permutation of coordinates of $X$ has on the vector of entropies $H(\Lambda) = (h_I(\Lambda))_{I\subset[d]}$. □

### 2.3.1 Supermodularity of the entropy map

As it is the case for real Gaussians, we would like the vector of entropies $H(\Lambda) := (h_I(\Lambda))$ to have values in the supermodular cone $\mathcal{S}_d$. As a first step towards proving this result, notice that the previous lemma implies that, if $\Lambda$ is a lattice such that $H(\Lambda) \in \mathcal{S}_d$, then for any diagonal matrix $D_a$ we still have $H(D_a\Lambda) \in \mathcal{S}_d$. Also for any permutation $\sigma$ of $\{1, \dots, d\}$, we still have $H(P^\sigma\Lambda) \in \mathcal{S}_d$.

In this subsection, we assume that $K$ is a local field and fix a uniformizer $\varpi$ of $K$ and a set $\mathcal{T} \subset \mathcal{O}$ of representative of the residue field $k$ (as in Proposition 1.15).

**Definition 2.11.** We say that a matrix $A = (a_{ij})_{1\le i,j\le d} \in \mathrm{GL}(d, K)$ is in *Hermite normal form* if it satisfies the following conditions:

1. $a_{ij} = 0$ for all $1 \le i < j \le d$; that is $A$ is lower triangular.
2. $a_{ii} = \varpi^{n_i}$ with $n_i \in \mathbb{Z}$ for all $1 \le i \le d$.
3. $a_{ij}$ is either 0 or is Laurent polynomial in $\varpi$ with coefficients in $\mathcal{T}$ of degree strictly less than $n_i$ for any $1 \le j < i \le d$.

**Lemma 2.12.** *For any matrix $A \in \mathrm{GL}(d, K)$ there exists a unique matrix $H$ in Hermite normal form and a unique matrix $U \in \mathrm{GL}(d, \mathcal{O})$ such that*

$$A = HU.$$

*In particular, any lattice $L = A\mathcal{O}^d \subset K^d$ has a unique representative $H \in \mathrm{GL}(d, K)$ in Hermite normal form.*

We can now state the second result of this section concerning the supermodularity of the entropy map. However, we first need to give an equivalent definition of the supermodular cone as

$$\mathcal{S}_d = \left\{(x_I)_{I\subset[d]} \in \mathbb{R}^{2^d} \colon x_\emptyset = 0 \text{ and } x_{Ii} + x_{Ij} \le x_I + x_{Iij}, \text{ for any } I \subset [d], i \ne j \in [d] \setminus I\right\}$$

where we write $Ii$ instead of $I \cup \{i\}$. These are the facet-defining inequalities of the cone $\mathcal{S}_d$ and there are $d(d-1)2^{d-3}$ of them; see [104].

**Theorem 2.13.** *The image of the map $H : \Lambda \to (h_I(\Lambda))_{I\subset[d]}$ lies in the supermodular cone $\mathcal{S}_d$, i.e, for any subset $I \subset [d]$ with $|I| \le d-2$ and $i \ne j \in [d] \setminus I$,*

$$h_{Ii}(\Lambda) + h_{Ij}(\Lambda) \le h_I(\Lambda) + h_{Iij}(\Lambda).$$

*Proof.* We prove this by induction on $d$. The result is trivial for $d = 1, 2$. Assume that it holds for lattices in $K^r$ for any $r \le d$, where $d \ge 3$. Let $\Lambda$ be a lattice in $K^d$ and $A$ its Hermite normal form. For any $I \subset [d]$ of size $|I| < d-2$ the inequality $h_{Ii}(\Lambda)+h_{Ij}(\Lambda) \le h_I(\Lambda)+h_{Iij}(\Lambda)$ holds for any $i \ne j$ not in $I$, thanks to the induction hypothesis. This is because, when

$|I| \leq d-2$, we are working on the lattice $\Lambda_{Ij}$ which is a lattice in dimension less than $d$. Then, it suffices to show the inequality when $I$ has size $d-2$. By Lemma 2.10 we can assume that $I = \{1, \dots, d-2\}$ and $i = d-1$ and $j = d$ (if not, we can just act on $\Lambda$ by a suitable permutation matrix). Let us write down the matrix $A$ as follows

$$A = \begin{pmatrix} \varpi^{a_1} & 0 & \dots & 0 & 0 & 0 \\ * & \varpi^{a_2} & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 & 0 \\ * & \dots & * & \varpi^{a_{d-2}} & 0 & 0 \\ * & \dots & * & * & \varpi^{a_{d-1}} & 0 \\ * & \dots & * & * & x & \varpi^{a_d} \end{pmatrix}.$$

Recall that since $A$ is the Hermite form of $\Lambda$ we have $\mathrm{val}(x) < a_d$ or $x = 0$. Now we have

$$h_{Ii}(\Lambda) = a_1 + \dots + a_{d-1}, \qquad h_{Ij}(\Lambda) = a_1 + \dots + a_{d-2} + \min(\mathrm{val}(x), a_d)$$
$$h_I(\Lambda) = a_1 + \dots + a_{d-2}, \quad \text{and} \quad h_{Iij}(\Lambda) = a_1 + \dots + a_d.$$

The inequality $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) \leq h_I(\Lambda) + h_{Iij}(\Lambda)$ then holds simply because $\mathrm{val}(x) \leq a_d$ and this finishes the proof. □

This theorem underlines another similarity between the local field Gaussians defined in [65] and classical multivariate Gaussian measures. From Lemma 2.10 we can see that acting on $\Lambda$ by a diagonal matrix just moves the point $H(\Lambda) \in \mathcal{S}_d$ in parallel to the *lineality* space of the cone $\mathcal{S}_d$, that is, the biggest vector space contained in $\mathcal{S}_d$.

The classical entropy map is tightly related to conditional independence. More precisely, if $\Sigma \in \mathrm{PD}_d$ and $X$ is a Gaussian vector with covariance matrix $\Sigma$, then for any $I \subset [d]$ and $i \neq j$ not in $I$ the variables $X_i$ and $X_j$ are independent given the vector $X_I$ if and only if $h_{Ii}(\Sigma) + h_{Ij}(\Sigma) = h_I(\Sigma) + h_{Iij}(\Sigma)$ and we write

$$X_i \perp\!\!\!\perp X_j | X_I \iff h_{Ii}(\Sigma) + h_{Ij}(\Sigma) = h_I(\Sigma) + h_{Iij}(\Sigma).$$

This means that the conditional independence models are exactly the inverse images by $H$ of the faces of $\mathcal{S}_d$ [155, Proposition 4.1]. It turns out that, in the local field setting, the non-archimedian entropy map $H$ defined in (2.3) also encodes conditional independence information on the coordinates of the random Gaussian vector $X$ as stated in the following proposition.

**Proposition 2.14.** *Assume $d \geq 2$, let $I$ be a subset of $[d]$ and let $i \neq j \in [d] \setminus I$ two distinct integers. Let $\Lambda$ be a lattice in $K^d$ and $X$ a random Gaussian vector with distribution given by $\Lambda$. Then the conditional independence statement $X_i \perp\!\!\!\perp X_j | X_I$ holds if and only if $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) = h_I(\Lambda) + h_{Iij}(\Lambda)$.*

*Proof.* Using Lemma 2.10 we reduce to the case $I = [r]$ where $r \leq d-2$ , $i = r+1$ and $j = i+1$. Let $A = (a_{i,j})$ be the unique representative in Hermite form of $\Lambda$. We claim

that $X_i \perp\!\!\!\perp X_j | X_I$ if and only if $a_{j,i} = 0$. To see why, let $Z = A^{-1}X$ which is a Gaussian vector whose distribution is the uniform on $\mathcal{O}^d$. We have $X_i = a_{i,1}Z_1 + \cdots + a_{i,i}Z_i$ and $X_j = a_{j,1}Z_1 + \cdots + a_{j,j}Z_j$. Since $Z_I = A_{I,I}^{-1}X_I$, given $X_I$ we know $Z_I$ and vice-versa. Hence $X_i \perp\!\!\!\perp X_j | X_I$ holds if and only if $(a_{j,i}Z_i + a_{j,j}Z_j) \perp\!\!\!\perp Z_i$. This happens if and only if the vectors $(1, 0)$ and $(a_{j,i}, a_{j,j})$ in $K^2$ are orthogonal; see [65]. This is equivalent to $\mathrm{val}(a_{j,j}) \leq \mathrm{val}(a_{j,i})$ which means that $a_{j,i} = 0$ since $A$ is in Hermite form. On the other hand, since $A$ is lower triangular, we have

$$h_I(\Lambda) = \mathrm{val}(\det(A_{I\times I}))\ ,\ \ h_{Ii}(\Lambda) = h_I(\Lambda) + \mathrm{val}(a_{i,i})$$
$$h_{Ij}(\Lambda) = h_I(\Lambda) + \min(\mathrm{val}(a_{j,i}), \mathrm{val}(a_{j,j})) \text{ and } h_{Iij}(\Lambda) = h_I(\Lambda) + \mathrm{val}(a_{i,i}) + \mathrm{val}(a_{j,j}).$$

So the equality $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) = h_I(\Lambda) + h_{Iij}(\Lambda)$ holds if and only if $\mathrm{val}(a_{j,j}) \leq \mathrm{val}(a_{j,i})$ since $A$ is the Hermite form of $\Lambda$ this happens if and only if $a_{j,i} = 0$. In combination with the calculation above, this finishes the proof. □

In other terms, the conditional independence statement $X_i \perp\!\!\!\perp X_j | X_I$ holds if and only if the entropy vector $H(\Lambda) = (h_I(\Lambda))$ is on the face of the polyhedral cone $\mathcal{S}_d$ cut by the equation $h_{Ii}(\Lambda) + h_{Ij}(\Lambda) = h_I(\Lambda) + h_{Iij}(\Lambda)$. This gives an analog of [155, Proposition 4.1].

**Corollary 2.15.** *The Gaussian conditional independence models are exactly those subsets of lattices that arise as inverse images of the faces of $\mathcal{S}_d$ under the map $H$.*

*Proof.* This follows immediately from the previous proposition. □

This underlines the importance of the map $H$, and also gives reason to think that the suitable analog of the positive definite cone on local fields is the set of lattices or more precisely the Bruhat-Tits building [1, 54]. A hard question in information theory for classical multivariate Gaussians is to describe the image of the entropy map [155]. This problem turns out to be difficult in our non-archimedean setting as well.

**Problem 2.3.2.** Characterize the image of the entropy map $H$ and describe how it intersects the faces of the supermodular cone $\mathcal{S}_d$. What can you say about the fibers of this map?

**Remark 2.16.** We recall that for any $d \geq 1$ the image $\mathrm{im}(H)$ is invariant under the action of the symmetric group and by translation in parallel to the lineality space of $\mathcal{S}_d$. This is thanks to Lemma 2.10. We will provide an answer for Problem 2.3.2 when $d = 2, 3$ at the end of Section 2.4.

We now provide an algorithm to compute the entropy vector $H(\Lambda)$, i.e, the coefficients of the polynomial $\varphi_\Lambda$. This relies on computing the Hermite form rather than directly solving the optimization problems given by equation (2.6).

**Algorithm 1:** Computing $H(\Lambda)$

**Input:** A full rank matrix $A = (a_1, \dots, a_n) \in K^{d\times n}$ with $n \geq d$ generating $\Lambda$
**Output:** The entropy vector $H(\Lambda)$
**for** $I \subset [d]$ **do**
  Compute the Hermite form $A_I$ of $\Lambda_I$.
  $h_I(\Lambda) \leftarrow \mathrm{val}(\det(A_I))$ (sum of valuations of diagonal elements of $A_I$)
**end**
$H(\Lambda) \leftarrow (h_I(\Lambda))_{I\subset[d]}$
**return** $H(\Lambda)$.

Let us now discuss a couple of low-dimensional examples when $K = \mathbb{Q}_p$.

**Example 2.17.** Let $\Lambda$ be the lattice represented by $A = \begin{pmatrix} 1 & 0 \\ p & p^2 \end{pmatrix}$. The coefficients $h_I(\Lambda)$ of the polynomial $\varphi_\Lambda$ can be computed from the representative $A$ using Line 1 and we have

$$h_\emptyset(\Lambda) = 0, \quad h_1(\Lambda) = 0, \quad h_2(\Lambda) = 1, \quad h_{1,2}(\Lambda) = 2$$

and then we get

$$\varphi_\Lambda(v_1, v_2) = \max(0, \quad v_1, \quad v_2 - 1, \quad v_1 + v_2 - 2).$$

The independence statement $X_1 \perp\!\!\!\perp X_2$ does not hold since the inequality $h_1(\Lambda) + h_2(\Lambda) \leq h_{12}(\Lambda)$ is strict. The tropical curve of $\varphi_\Lambda$ and its regular triangulation of the square are shown in Example 2.17.

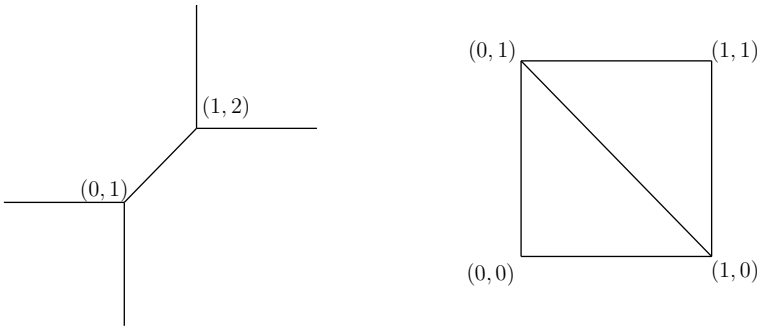



Figure 2.1: Tropical curve of $\varphi_\Lambda$ and its regular triangulation of the square for Example 2.17

**Example 2.18.** Let $\Lambda$ be the lattice represented by $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \varpi^2 & 0 \\ 1 & \varpi & \varpi^2 \end{pmatrix}$. The polynomial

$\varphi_\Lambda$ can be computed again using Algorithm 1 and we get

$$\begin{array}{ccc} & h_\emptyset(\Lambda)=0 & \\ h_1(\Lambda)=0, & h_2(\Lambda)=0, & h_3(\Lambda)=0 \\ h_{1,2}(\Lambda)=2, & h_{1,3}(\Lambda)=1, & h_{2,3}(\Lambda)=1 \\ & h_{1,2,3}(\Lambda)=4. & \end{array}$$

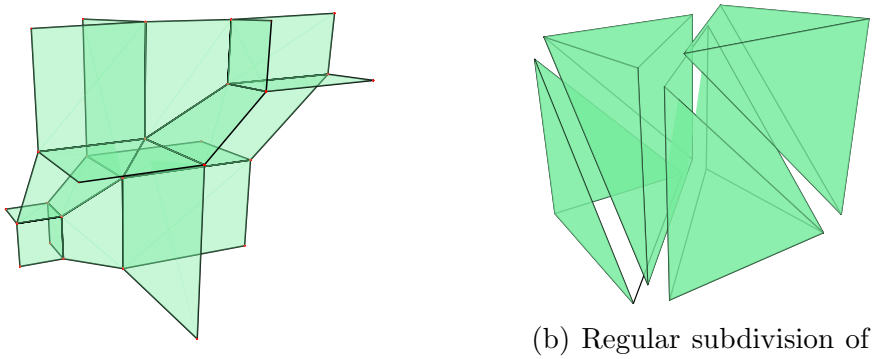
So we deduce that

$$\varphi_\Lambda(v)=\max(0,v_1,v_2,v_3,v_1+v_2-2,v_1+v_3-1,v_2+v_3-1,v_1+v_2+v_3-4).$$

We can easily check that the supermodularity inequalities are satisfied. Also, none of the conditional independence statements $X_i \perp\!\!\!\perp X_j|X_k$ are satisfied for $\{i,j,k\}=\{1,2,3\}$ since the point $H(\Lambda)$ is in the interior of the cone $\mathcal{S}_3$, i.e, all the inequalities

$$h_{ki}(\Lambda)+h_{kj}(\Lambda)\leq h_i(\Lambda)+h_{ijk}(\Lambda)$$

are strict. Figure 2.2 shows the tropical geometry for the lattice $\Lambda$.



(a) Tropical variety of $\varphi_\Lambda$.

(b) Regular subdivision of the Newton polytope of $\varphi_\Lambda$.

Figure 2.2: Tropical geometry of the lattice $\Lambda$ for Example 2.18.

**Remark 2.19.** For any lattice $\Lambda$, there exists a maximal (for inclusion) diagonal lattice inside $\Lambda$ and a minimal diagonal lattice containing $\Lambda$. Let us denote these two lattices by $\boldsymbol{\varpi}^{\boldsymbol{a}}$ and $\boldsymbol{\varpi}^{\boldsymbol{b}}$ respectively, where $a\geq b\in\mathbb{Z}^d$. So, we have the inclusions $\boldsymbol{\varpi}^{\boldsymbol{a}}\subset\Lambda\subset\boldsymbol{\varpi}^{\boldsymbol{b}}$. It is not difficult to see that the region of linearity corresponding to the monomial $v_1+\cdots+v_d-h(\Lambda)$ in the tropical polynomial $\varphi_\Lambda(v)$ is the orthant $\mathbb{R}_{\geq a}:=\{x\in\mathbb{R}^d, x\geq a\}$. Similarly, the region of linearity corresponding to the monomial 0 is the orthant $\mathbb{R}_{\leq b}:=\{x\in\mathbb{R}^d, x\leq b\}$. From this, we can the deduce the recursive relation

$$h_{[d]}(\Lambda)=h_{[d-1]}(\Lambda)+a_d.$$

This iterative way of computing the entropy map $H(\Lambda)$ is slightly more efficient than Line 1 where we have to compute the whole Hermite form of $\Lambda_I$ for every $I\subset[d]$. We provide an implementation of Algorithm 1 in the repository (2.4).

## 2.4 The entropy map on nonarchimedean fields

In this section we generalize some of the results in Section 2.3 to the case where $K$ is a field with a nonarchimedean valuation. When the residue field $k$ of $K$ is infinite or the valuation group $\Gamma$ is dense in $\mathbb{R}$, the probabilistic framework we had in Section 2.3 is no longer valid. More precisely, we lose the local compactness and we no longer necessarily have a Haar measure on $K$.

We define the entropy $H(\Lambda) = (h_I(\Lambda))_{I\subset[d]}$ of a lattice as in Section 2.3, i.e for any subset $I \subset [d]$,

$$h_I(\Lambda) \coloneqq \min_{|J|=|I|} \mathrm{val}(\det(A_{I\times J})),$$

where $A$ is a representative of $\Lambda$. We can still define a *Hermite representative* of $\Lambda$.

The argument used in Theorem 2.13 can be used again to show that the image of $H$ still lies in the supermodular cone $\mathcal{S}_d$. In this setting however, since the valuation group can be dense in $\mathbb{R}$, the image is not necessarily in $\mathcal{S}_d \cap \mathbb{Z}^{2^d-1}$. As in Section 2.3, the map $H$ fails to be surjective when $d \geq 3$. The algorithm we provide in (2.4) computes the map $H$ when $K = \mathbb{Q}\{\{t\}\}$ is the field of Puiseux series over $\mathbb{Q}$.

Now we show that the only distribution on the field Laurent series $K = \mathbb{R}((t))$ that satisfies the definition suggested in [65, Definition 4.1] is the Dirac measure at 0. Let $\mathbb{P}$ be such a probability measure. First, we recall that if $X$ is a random variable with distribution $\mathbb{P}$, then for any $a \in \mathcal{O}_{\mathbb{K}}^{\times}$ the random variables $X$ and $aX$ have the same distribution, and we write $X \overset{d}{=} aX$. In particular, for any $a \in \mathbb{R}^{\times}$ we have $X \overset{d}{=} aX$.

**Proposition 2.20.** *The probability distribution $\mathbb{P}$ is the Dirac measure at 0.*

*Proof.* We can write the power series expansion of $X$ as $X = X_0t^V + X_1t^{V+1} + \dots$, where $V \in \mathbb{Z}$ is the random valuation of $X$. Hence for $a \in \mathbb{R}^{\times}$ we have $aX = aX_0t^V + aX_1t^{V+1} + \dots$, and we deduce that $X_k \overset{d}{=} aX_k$ for any $k \geq 0$ and $a \in \mathbb{R}^{\times}$. We then deduce that $X_k = 0$ almost surely for all $k \geq 0$. Hence $X = 0$ almost surely which finishes the proof. □

Using a variant of this argument, it is not difficult to see that a similar problem would arise when we try to define Gaussian measures by orthogonality for all fields listed in Example 2.3. It is not immediately clear how to fix this problem and find a suitable definition for *Gaussian measures* on nonarchimedean valued fields.

**Problem 2.4.1.** Is there a suitable definition for Gaussian measures on the fields listed Example 2.3?

**Remark 2.21.** We can define a probability measure on $\mathbb{R}^d$ induced by $\Lambda$ via its tail distribution $Q_\Lambda$ as in Section 2.3. One can see that the support of this distribution is $\mathrm{trop}(\Lambda) \coloneqq \mathrm{val}(\Lambda \cap (K^{\times})^d)$; the image under valuation of points in $\Lambda$ with no zero coordinates. This is in general a polyhedral complex in $\mathbb{R}^d$ where each edge is parallel to some $e_I \coloneqq \sum_{i\in I} e_i$. The following figure is a drawing of $\mathrm{trop}(\Lambda)$ for a lattice in $K^3$ when $K = \mathbb{K}$ (the field of generalized Puiseux series).

Figure 2.3 is a depiction of $\mathrm{trop}(\Lambda)$ for the lattice $\Lambda$ in Example 2.18. In this case, $\mathrm{trop}(\Lambda)$ is a polyhedral complex is composed of a bounded line segment attached to three unbounded 2-dimensional polyhedra. These polyhedra are in turn attached to an orthant of dimension 3.

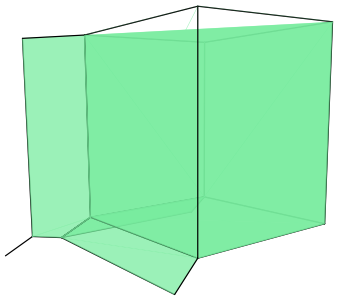


Figure 2.3: The polyhedral complex $\mathrm{trop}(\Lambda)$ for $\Lambda$ in Example 2.18.

To conclude this section we give a partial answer for Problem 2.3.2 when $d = 2, 3$ and the valuation group is $\mathbb{R}$.

**Proposition 2.22.** *For $d = 2$, the image $\mathrm{im}(H)$ of the entropy map $H$ is exactly $\mathcal{S}_2$.*

*Proof.* For $\Lambda$ with representative $\begin{pmatrix} t^a & 0 \\ t^b & t^{b+\delta} \end{pmatrix}$ with $a, b \in \mathbb{R}$ and $\delta \geq 0$ we have $H(\Lambda) = (a, b, a + b + \delta)$. So $H$ is indeed surjective onto $\mathcal{S}_2$. □

For $d = 3$, the cone $\mathcal{S}_3 \subset \mathbb{R}^7$ has a lineality space $\mathcal{L}_3$ of dimension 3. Since both $\mathcal{S}_3$ and $\mathrm{im}(H)$ are stable under translations in $\mathcal{L}_3$ (see Remark 2.16 and Lemma 2.10 on diagonal scaling of lattices), they are fully determined by their projection onto a complement of $\mathcal{L}_3$. Let us we write vectors $x$ of $\mathbb{R}^7$ in the form

$$x = (x_1, x_2, x_3;\ x_{12}, x_{13}, x_{23};\ x_{123}),$$

and let us project $\mathcal{S}_3$ and $\mathrm{im}(H)$ on the linear space $W \subset \mathbb{R}^7$ of vectors of the form

$$x = (0, x_2, x_3; 0, x_{13}, x_{23}; 0).$$

which is a complement of $\mathcal{L}_3$ in $\mathbb{R}^7$. We write a vector of $W$ as $(x_2, x_3; x_{13}, x_{23})$ or simply as $(w, x, y, z)$ to simplify notation. Let us denote by $\mathcal{P}, \mathcal{C}$ the projections of $\mathrm{im}(H)$ and $\mathcal{S}_3$ respectively onto the space $W$. From Section 2.3.1, we clearly have $\mathcal{P} \subset \mathcal{C}$.

The projection $\mathcal{C}$ of $\mathcal{S}_3$ onto $W$ is a polyhedral cone that does not contains any lines. In the language of polyhedral geometry, this is called a *pointed cone*. Moreover, the dimension of this projection is 4. It is defined in $W$ by the inequalities

$$\begin{cases} w \leq 0, & x \leq y, \\ w + x \leq z, & y \leq 0, \\ z \leq w, & y + z \leq x. \end{cases} \tag{2.9}$$

This defines $\mathcal{C}$ as a pointed cone over a bipyramid (see Figure 2.4).

On the other hand, any lattice $\Lambda$ in $\mathbb{K}^3$ can be represented, up to diagonal scaling, by a representative with Hermite form of the shape

$$\begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix}.$$

The entropy vector of a lattice $\Lambda$ with such a Hermite normal form is of the shape

$$H(\Lambda) = (0, h_2, h_3; 0, h_{13}, h_{23}; 0).$$

This corresponds to the projection of $\mathrm{im}(H)$ to $W$ parallel to $\mathcal{L}_3$. So the projection $\mathcal{P}$ of $\mathrm{im}(H)$ onto $W$ is the set

$$\mathcal{P} = \left\{ H(\Lambda),\ \Lambda \text{ given by a matrix of the shape } \begin{pmatrix} 1 & 0 & 0 \\ * & 1 & 0 \\ * & * & 1 \end{pmatrix} \text{ in } \mathbb{K}^{3\times 3} \right\}.$$

For a lattice $\Lambda$ with representative $A = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}$, such that $a, b, c \in \mathbb{K}$ with negative or zero valuation (see Definition 2.11), the point $H(\Lambda)$ in $W$ is given by

$$\begin{cases} w = h_2(\Lambda) = \mathrm{val}(a), \\ x = h_3(\Lambda) = \min(\mathrm{val}(b), \mathrm{val}(c)), \\ y = h_{13}(\Lambda) = \mathrm{val}(c), \\ z = h_{23}(\Lambda) = \min(\mathrm{val}(ac - b), \mathrm{val}(a)). \end{cases}$$

One can check that, for any choice of $a, b, c \in \mathbb{K}$ with negative or zero valuation, the above coordinates satisfy the inequalities in (2.9). With the constraints on the valuations of $a, b, c$, and from this parametric representation of $\mathcal{P}$, we can see that points of $\mathcal{P}$ have to satisfy the inequalities

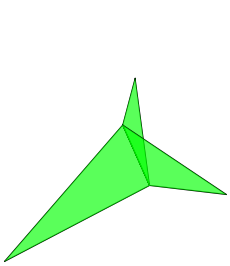
$$\begin{cases} w \leq 0, \\ x \leq y, \\ y \leq 0. \end{cases}$$

The only part that remains to determine is the inequalities involving the last variable $z$. The ambiguity comes from the fact that cancellations can happen in $ac - b$ which might

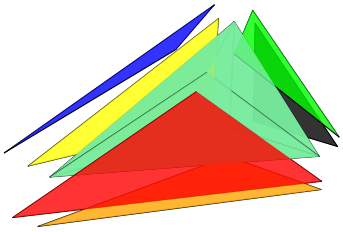
affect $\mathrm{val}(ac-b)$ and hence also $z$. But, separating the cases where $\mathrm{val}(ac) = \mathrm{val}(b)$ and $\mathrm{val}(ac) \neq \mathrm{val}(b)$, we get the following three sets of inequalities that describe $\mathcal{P}$ as a polyhedral complex:

$$\begin{cases} w \leq 0, \\ x \leq w+y, \\ y \leq 0, \\ z = x, \end{cases} \quad , \quad \begin{cases} w \leq 0, \\ x \leq y, \\ y \leq 0, \\ y+w \leq x, \\ z = y+w, \end{cases} \quad \text{and} \quad \begin{cases} w \leq 0, \\ y \leq 0, \\ x = y+w, \\ z \leq w, \\ x \leq z. \end{cases}$$

We can then see that $\mathcal{P}$ is a polyhedral fan of dimension 3 inside $\mathcal{C}$. More precisely, $\mathcal{P}$ is the union of three pointed polyhedral cones of dimension 3 inside $\mathcal{C}$ which is a cone of dimension 4. Figure 2.4 depicts the intersections of $\mathcal{P}$ and $\mathcal{C}$ with the hyperplane $w+x+y+z+1=0$ (slicing the pointed cones with a hyperplane).



(a) $\mathcal{P} \cap \{w+x+y+z+1=0\}$.



(b) $\mathcal{C} \cap \{w+x+y+z+1=0\}$. Red facet: $1 \perp\!\!\!\perp 2$; Blue facet: $1 \perp\!\!\!\perp 3$, Green facet: $2 \perp\!\!\!\perp 3$; Orange facet: $1 \perp\!\!\!\perp 2|3$; Yellow facet: $1 \perp\!\!\!\perp 3|2$; Grey facet: $2 \perp\!\!\!\perp 3|1$.

Figure 2.4: Intersections of $\mathcal{P}$ and $\mathcal{C}$ with the affine hyperplane $x+y+z+w+1=0$.

**Corollary 2.23.** *The entropy map* $H : \mathrm{GL}(d,\mathbb{K})/\mathrm{GL}(d,\mathcal{O}_{\mathbb{K}}) \to \mathcal{S}_d$ *is not surjective when* $d \geq 3$.

We expect this result to hold in every dimension; that is the image $\mathrm{im}(H)$ is a polyhedral fan whose facets are polyhedral cones of dimension $\frac{d(d+1)}{2}$ inside $\mathcal{S}_d$ which is of dimension $2^d - 1$.

## 2.5 Statistical models in the Bruhat-Tits building

Since the Bruhat-Tits building is the parameter space of non-degenerate Gaussian distribution on $K^d$, Gaussian statistical models are subsets of this building. Describing what certain models look like in this setting turns out to be a hard task that requires more advanced techniques. In this section we give a couple of examples of such models.

In statistics, given a data set of points, one usually tries to the best probability distribution (in a certain model) that fits the data. This fitting is often done by maximizing the likelihood function on a certain model of distributions. The following theorem addresses the case where we fit a Gaussian distribution to a collection of data points; that is our model is the entire set of lattices.

**Theorem 2.24.** *Let $\mathcal{X} = \{x_1, \dots, x_N\}$ be a dataset of points in $K^d$ of full rank. Then there is a unique non-degenerate Gaussian distribution $\mathbb{P}$ on $K^d$ that maximizes the likelihood*

$$\mathcal{L}(\mathcal{X}, L) = \prod_{x \in \mathcal{X}} \frac{1[x \in L]}{\mu^{\otimes d}(L)},$$

*Namely this distribution is the Gaussian distribution corresponding to the lattice*

$$L_{\mathcal{X}} = \mathrm{span}_{\mathcal{O}}(\mathcal{X}).$$

*Proof.* Define the $\mathcal{O}$-module $L_{\mathcal{X}} := \mathrm{span}_{\mathcal{O}}(\mathcal{X})$. Since $\mathcal{X}$ is of full rank, $L_{\mathcal{X}}$ is a lattice in $K^d$. Now let $L$ be any other lattice that contains $\mathcal{X}$. Then $\mathcal{L}(\mathcal{X}, L) = \mu^{\otimes d}(L)^{-1}$ and $\mathcal{L}(\mathcal{X}, L_{\mathcal{X}}) = \mu^{\otimes d}(L_{\mathcal{X}})^{-1}$. Since $L_{\mathcal{X}} \subset L$ we have $\mu^{\otimes d}(L) \geq \mu^{\otimes d}(L_{\mathcal{X}})$. Thus $\mathcal{L}(\mathcal{X}, L) \leq \mathcal{L}(\mathcal{X}, L_{\mathcal{X}})$. The lattice $L_{\mathcal{X}}$ maximizes the likelihood. Suppose that $L_{\mathcal{X}} \subsetneq L$ by means of a basis change without loss of generality we can suppose that $L = \mathcal{O}^d$. There exists an orthogonal matrix $U$ and a diagonal matrix $D := \mathrm{diag}(\varpi^{n_1}, \dots, \varpi^{n_d})$ such that $L_{\mathcal{X}} = UD\mathcal{O}^d$. Since $L_{\mathcal{X}} \subsetneq L$ we have $n_i \geq 0$ for any $1 \leq i \leq d$ and there exists $k$ such that $n_k > 0$. Thus $\mu^{\otimes d}(L_{\mathcal{X}}) < \mu^{\otimes d}(L)$. Then $L_{\mathcal{X}}$ is the unique lattice that maximizes the likelihood. □

**Remark 2.25.** In the case where $\mathcal{X}$ spans a proper subspace $W_{\mathcal{X}} := \mathrm{span}_K(\mathcal{X})$ of $K^d$ we can define a Haar measure $\lambda$ on $W_{\mathcal{X}}$ and the likelihood function defined for every full rank lattice of $W_{\mathcal{X}}$ as $\mathcal{L}(\mathcal{X}, L) = \prod_{x \in \mathcal{X}} 1_{x_i \in L} / \mu^{\otimes d}(L)$. The maximum likelihood estimate in this case is again $L_{\mathcal{X}} := \mathrm{span}_{\mathcal{O}_K}(\mathcal{X})$, and it is the minimal lattice with respect to inclusion amongst those that maximize the likelihood.

### 2.5.1 Conditional independence models

**Theorem 2.26.** *Let $K$ be a local field and $q$ the cardinality of its residue field $k \cong \mathbb{F}_q$. Let $X := (X_1, \dots, X_d)^T$ be a Gaussian vector in $K^d$ and $I$ a proper subset of $[d]$. The maximal subsets $J := \{j_1, \dots, j_r\}$ of $\{1, \dots, d\} \setminus I$ such that $X_{j_1}, \dots X_{j_r}$ are mutually independent given $X_I$ are the bases of an $\mathbb{F}_q$-realizable matroid with base set $\{1, \dots, d\} \setminus I$ where $q$ is the size of the residue field $k$.*

*Proof of Theorem 2.26.* Without loss of generality we can suppose that $I = \{1, \dots, \ell\}$ for some $\ell < d$. By Lemma 2.12, there exists a matrix $A$ in Hermite normal form such that the support of $X$ is the lattice $L = A.\mathcal{O}_K^d$. Then there exist a standard Gaussian vector $Y$ such that $X = AY$. Let $B = (A_{ij})_{\ell+1 \leq i,j \leq d}$ be the lower-right $(d-\ell) \times (d-\ell)$ block of $A$, and

$\{f_i : \ell+1 \leq i \leq d\}$ the linear forms defined by the rows of $B$. For a subset $J := \{j_1, \dots, j_r\}$ of $1, \dots, d \setminus I$, by Lemma 1.30, we have that $X_{j_1}, \dots, X_{j_r}$ are mutually independent given $X_I$ if and only if $\{f_{j_i} : 1 \leq i \leq r\}$ are orthogonal. Let us define the matrix $C$ as the matrix obtained from $B \in \mathrm{GL}(d-\ell, K)$ by scaling its rows so that they have norm 1. Then from Proposition 1.20 we deduce that $X_{j_1}, \dots, X_{j_r}$ are independent given $X_I$ if and only if the images (modulo $\varpi$) of the rows of $C$ indexed by $J$ are linearly independent in $k^{d-\ell}$. So conditional independence given $X_I$ is encoded in an $k$-realizable matroid given by the images modulo $\varpi$ of the rows of $C$. □

**Example 2.27.** Let $p$ be a prime number and consider the Gaussian vector $X = (X_i)_{1\leq i\leq 4}$ in $\mathbb{Q}_p^4$ with distribution given by the following lattice

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & p & 0 \\ 1 & p^{-1} & p^{-1} & p^2 \end{bmatrix} \mathbb{Z}_p^4.$$

If $I = \{1\}$ then the matrices $B$ and $C$ from the proof of Theorem 2.26 in this case are

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & p & 0 \\ p^{-1} & p^{-1} & p^2 \end{bmatrix} \in \mathbb{Z}_p^{3\times 3} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \in \mathbb{F}_p^{3\times 3}.$$

Then from the matrix $C$ we deduce the following conditional independence statements

$$X_2 \perp\!\!\!\perp X_4 | X_1, \quad X_3 \perp\!\!\!\perp X_4 | X_1 \quad \text{and} \quad X_3 \perp\!\!\!\perp X_4 | X_1,$$

However the three variables $X_2, X_3$ and $X_4$ are not mutually independent conditioned $X_1$. Independence conditioned on $X_1$ is encoded in the matroid that arises from the rows of $C$.

Let $I = \{1, \dots, r\}$ and $J = \{r+1, \dots, r+s\}$ where $r, s \geq 0$ are integers with $r + s \leq d$. We denote by $\mathcal{M}_{I,J}$ the model of non-degenerate Gaussian distributions on $\mathbb{Q}_p^d$ such that the variables indexed by $J$ are all independent given those indexed by $I$ i.e. the set of rank $d$ lattices $L$ in $K^d$ such that if $X$ is a Gaussian on with lattice $L$ we get

$$X_{r+1} \perp\!\!\!\perp \dots \perp\!\!\!\perp X_{r+s} \mid X_1, \dots, X_r. \tag{2.10}$$

Using Lemma 2.12 and Theorem 1.30 we can see that the points (or distributions) in $\mathcal{M}_{I,J}$

are exactly those lattices $L$ whose Hermite normal form has the following shape

$$\begin{bmatrix} * & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & & & & & & & \vdots \\ * & \dots & * & \ddots & & & & & & \vdots \\ * & \dots & * & * & \ddots & & & & & \vdots \\ * & \dots & * & 0 & * & \ddots & & & & \vdots \\ \vdots & & \vdots & \vdots & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \vdots & 0 & \dots & 0 & * & \ddots & & \vdots \\ * & \dots & * & * & \dots & * & * & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & 0 \\ * & \dots & * & * & \dots & * & * & \dots & * & * \end{bmatrix},$$

where the red block is the below diagonal part of the block indexed by $J \times J$. The situation is similar to Gaussians on the real numbers where conditional independence implies that certain entries of the concentration matrix have to be zero.

The reader might be wary of the particular choice of $I$ and $J$, but by permuting the variables $X_1, \dots, X_d$ we may always assume that $I, J$ have the prescribed shapes. Also, since the conditional independence statement (2.10) does not change when scaling the lattice $L$ it is natural to want to describe the model $\mathcal{M}_{I,J}$ in the Bruhat-Tits building.

**Problem 2.28.** What does the model $\mathcal{M}_{I,J}$ look like in the Bruhat-Tits building? What happens when we have multiple conditional independence statements? In the spirit of Theorem 2.24, can we easily fit conditional independence models to a collection of data points?

### 2.5.2 Exchangeable Gaussian vectors

Let $d \geq 1$ be an integer. We say that a probability distribution in $\mathbb{R}^n$ is exchangeable if it is invariant under the action of the symmetric group $S_d$ on $\mathbb{R}^n$. The centered Gaussian distributions in $\mathbb{R}^d$ which are exchangeable are exactly those distributions whose positive semidefinite covariance matrix $\Sigma = (\Sigma_{ij})$ satisfies

$$\Sigma_{11} = \cdots = \Sigma_{dd} \quad \text{and} \quad \Sigma_{ij} = \Sigma_{12} \quad \text{for any } 1 \leq i \neq j \leq d.$$

This gives a description of the set (or model) of exchangeable centered Gaussians in $\mathbb{R}^d$ as a subset of the positive semi-definite cone. A similar question can be formulated in the local field setting.

**Question 2.29.** When $K$ is a non-archimedean local field, what are the exchangeable Gaussian distributions on $K^d$? Equivalently what are the lattices in $K^d$ that are invariant under the action of $S_d$? What does this set of invariant lattices look like in the Bruhat-Tits Building?

This question falls under the general problem of determining the set of fixed points under the action of a (compact) group on the Bruhat-Tits building. We provide a partial answer in the case $d = 2$.

**Proposition 2.30.** *Let $L$ be a lattice in $K^2$. Then $L$ is invariant under the action of the symmetric group $S_2$ if and only if its Hermite normal form has the form*

$$\begin{pmatrix} \varpi^a & 0 \\ 0 & \varpi^a \end{pmatrix} \quad \text{with } a \in \mathbb{Z}.$$

*or*

$$\begin{pmatrix} \varpi^u & 0 \\ \varpi^u\zeta & \varpi^v \end{pmatrix} \quad \text{with } v > u \in \mathbb{Z} \text{ and } \zeta \in \mathcal{O}_K^\times \text{ such that } \zeta^2 = 1 \pmod{\varpi^{v-u}\mathcal{O}_K}.$$

*Proof.* Suppose that a lattice $L$ with Hermite normal form

$$\begin{pmatrix} \varpi^a & 0 \\ x & \varpi^b \end{pmatrix}$$

is invariant under the action of the involution $(1, 2)$. If $x = 0$ then it is clear that we must have $a = b$. Now suppose that $x \neq 0$, then we deduce that $L$ is also represented by the matrix

$$\begin{pmatrix} x & \varpi^b \\ \varpi^a & 0 \end{pmatrix}$$

hence also by the matrix

$$\begin{pmatrix} x & 0 \\ \varpi^a & \varpi^{b+a}/x \end{pmatrix}.$$

Write $x = \varpi^c\zeta$ with $\zeta \in \mathcal{O}_K^\times$ and $c < b \in \mathbb{Z}$ (by definition of Hermite normal form). Then $L$ is also represented by the matrix

$$\begin{pmatrix} \varpi^c & 0 \\ \varpi^a\zeta^{-1} & \varpi^{b-c+a} \end{pmatrix}.$$

Since the last matrix is also in Hermite normal form we deduce that

$$a = c \quad \text{and} \quad \zeta^{-1} = \zeta \pmod{\varpi^{b-a}}$$

Conversely, it is not hard to check that if $L$ is a lattice with one of the suggested Hermite normal forms then $L$ is invariant under action of the transposition $(1, 2)$ (which just swaps the rows of the representative matrices). □

**Example 2.31.** The following figure describes locally (i.e. we only drew a bounded region in the infinite 3-valent tree that is the Bruhat-Tits building $\mathcal{B}_2(\mathbb{Q}_2)$) the set of $S_2$-invariant lattice classes in the building $\mathcal{B}_2(\mathbb{Q}_2)$. This set consists of the lattice classes colored in blue

and red. The points colored in blue form an apartment (we only drew a piece of it) in the building $\mathcal{B}_2(\mathbb{Q}_2)$ while the red points are the points that have distance 1 from this apartment.

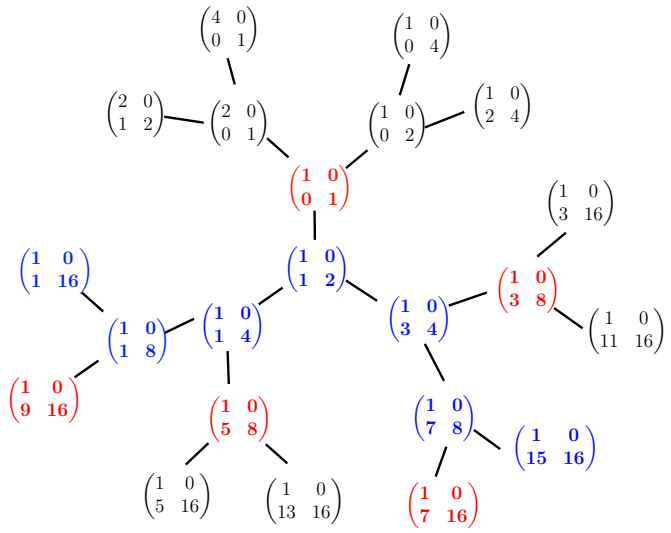


Figure 2.5: The set of $S_2$-invariant lattices in the building $\mathcal{B}_2(\mathbb{Q}_2)$ (colored in blue and red).

The following figure describes (locally) the set of $S_2$-invariant lattice classes in the building $\mathcal{B}_2(\mathbb{Q}_3)$ (i.e. exchangeable non-degenerate Gaussian distributions in $\mathbb{Q}_3^2$). This set consists of the lattice classes colored in blue.

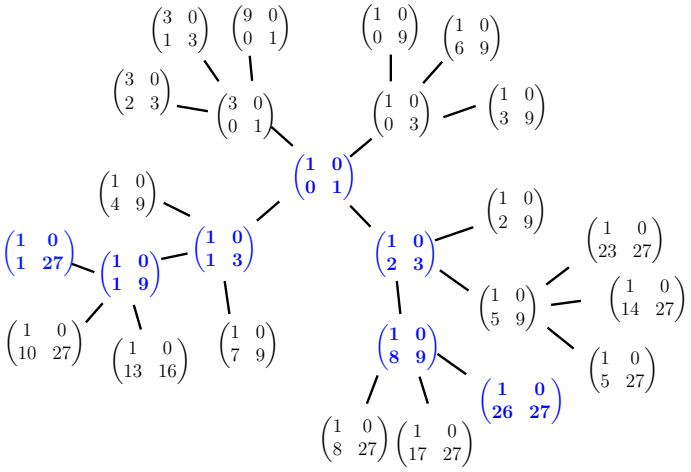


Figure 2.6: The set of $S_2$-invariant lattices in the building $\mathcal{B}_2(\mathbb{Q}_3)$ (colored in blue).

## 2.6 Conclusion

In conclusion, there are many similarities between the classical theory of Gaussian distributions on euclidean spaces and the theory of Gaussian measures on local fields as defined by Evans in [65]. In this chapter we have exhibited another similarity in terms of differential entropy, and studied a couple of Gaussian models in the Bruhat-Tits building. This gives reason to think that the suitable non-archimediean analog of the positive definite cone is indeed the set of lattices, or more precisely, in the language of group theorists, the Bruhat-Tits building for PGL. This analogy can still be carried out for non-archimedean valued fields in general. However, when the field $K$ has a dense valuation group or an infinite residue field, we lose the probabilistic interpretation and thus also the notion of entropy.

# Chapter 3

# Sampling from $p$-adic algebraic manifolds

This chapter is based on joint work [50] with Enis Kaya.

## 3.1 Introduction

Algebraic *varieties* and *manifolds* are ubiquitous in many areas of mathematics: In number theory, Shimura varieties arise as complex algebraic varieties that parameterize certain types of Hodge structures; Calabi-Yau manifolds model a number of phenomena in physics generally and superstring theory specifically, and enjoy interesting geometric properties; and many interesting probabilistic models arise as varieties whose algebraic and geometric properties have meaningful probabilistic and statistical interpretations.

While the use of $p$-adic numbers has not yet become as common a practice in many domains, they have started to find numerous applications, for example in mathematical physics [165]. Moreover, the interest and research activity addressing probabilistic and statistical questions in the $p$-adic setting have been gaining momentum starting from the early work of Evans [60, 64, 65], Bikulov, Vladimirov, Volovich and Zelenov [18, 165] to the more recent developments [16, 26, 28, 62, 105, 118] and Chapter 2 to mention a few . In this line of thought, it is quite desirable to have an efficient method of sampling from $p$-adic manifolds.

In this chapter, inspired by the work of Breiding and Marigliano [23], we tackle the problem of sampling from a $p$-adic algebraic variety with a prescribed probability distribution by intersecting it with random linear spaces of complementary dimension. Our results provide $p$-adic analogues of the results in [23] and are in line with $p$-adic analogs of *Crofton's formulas* in integral geometry (see [105] and [26] for $p$-adic integral geometry) which provide a link between the volume of a manifold and its expected number of intersection points with random linear spaces.

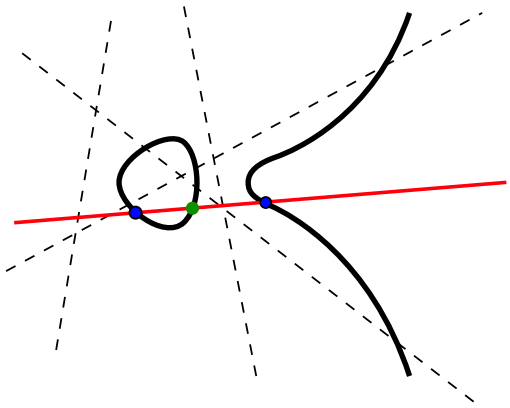


Figure 3.1: An illustration of the sampling method.

Over the real or complex numbers, sampling from manifolds (especially in the parameterized case) often involves using a Markov chain sampling (Metropolis algorithm, Gibbs sampler, hit-and-run algorithm etc.) [39, 9, 114]. While this method is fairly simple (both computationally and mathematically), it approximates the desired probability measure only asymptotically and requires a study of the mixing time of the Markov chain in question [40, 41]. Moreover, since nonarchimedean local fields are totally disconnected topological spaces, the usual Markov chain sampling algorithms are ill suited for such a setting. Our method has the advantage of sampling exactly from the desired probability density. Also, given its geometric nature, it works regardless of the nature of the topology involved.

To set things up, let $K$ be a nonarchimedean local field with a normalized discrete valuation map; $\mathrm{val}\colon K \twoheadrightarrow \mathbb{Z} \cup \{+\infty\}$. We fix once and for all a uniformizer $\varpi$; that is an element $\varpi \in K$ such that $\mathrm{val}(\varpi) = 1$. We denote by $\mathcal{O} := \{x \in K\colon \mathrm{val}(x) \geq 0\}$ the valuation ring of $K$; and by $k := \mathcal{O}/\varpi\mathcal{O}$ its residue field. It is known that $k$ is isomorphic to a finite field $\mathbb{F}_q$ with $q$ elements; where $q$ is a power of the characteristic $p := \mathrm{char}(k)$. The valuation val defines an absolute value $|\cdot|$ on $K$ by setting $|x| := q^{-\mathrm{val}(x)}$ for $x \in K$. We denote by $\mu$ the unique real-valued Haar measure on $K$ such that $\mu(\mathcal{O}) = 1$. We refer to Section 1.1.2 for an introduction to local fields and to [146, Part 1] for a detailed account.

**Remark 3.1.** nonarchimedean local fields come in two flavors. Those that have the same characteristic as their residue fields are isomorphic (as topological fields) to $\mathbb{F}_q((\varpi))$; the field of Laurent series in one variable $\varpi$ with coefficients in $\mathbb{F}_q$. The second type has characteristic 0 and are finite extensions of the field $\mathbb{Q}_p$ of $p$-adic numbers for some prime $p$.

Let us fix a positive integer $N \geq 2$. We denote by $\mathbb{A}^N = K^N$ the $N$-dimensional affine space over $K$. The space $\mathbb{A}^N$ inherits from $K$ the product Haar measure, which we denote by $\mu_{\mathbb{A}^N}$. When there is no risk of confusion, we simply denote this measure by $dx$. We endow

$\mathbb{A}^N$ with the norm

$$\|x\| = \max_{1\le i\le N} |x_i|, \quad \text{for } x = (x_1, \dots, x_N)^\top \in \mathbb{A}^N,$$

and the valuation

$$\mathrm{val}(x) = \min_{1\le i\le N} \mathrm{val}(x_i) = -\log_q(\|x\|), \quad \text{for } x = (x_1, \dots, x_N)^\top \in \mathbb{A}^N.$$

This makes $\mathbb{A}^N$ a metric space with the metric given by $d(x, y) = \|x - y\|$ for $x, y \in \mathbb{A}^N$. We refer the reader to Section 1.1.3 for more details on norms, valuation and orthogonality.

An affine algebraic variety in $\mathbb{A}^N$ is the zero set of a system of polynomials $\boldsymbol{p} = (p_1, \dots, p_r)$ in $K[x_1, \dots, x_N]$, i.e.

$$\{x \in \mathbb{A}^N : p_1(x) = \cdots = p_r(x) = 0\}.$$

We refer to smooth and irreducible[1] varieties over $K$ (affine or projective) as *algebraic manifolds*.

Although the notions of dimension and degree of an algebraic $K$-manifold $X$ are the usual notions from algebraic geometry, the notion of *volume* on $X$ is not as standard. Let $X \subset \mathbb{A}^N$ be an affine algebraic $K$-manifold of dimension $n$. For $\epsilon > 0$ and $x \in \mathbb{A}^N$, let us denote by $B_N(x, \epsilon) = \left\{y \in \mathbb{A}^N : d(x, y) \le \epsilon\right\}$ the ball of radius $\epsilon$ and center $x$. The *volume measure* $\mu_X$ on $X$ is defined as follows, for an open set in $V \subset X$:

$$\mu_X(V) \coloneqq \lim_{\epsilon\to 0} \frac{\mu_{\mathbb{A}^N}\left(\bigcup_{x\in V} B_N(x, \epsilon)\right)}{\mu_{\mathbb{A}^{N-n}}\left(B_{N-n}(0, \epsilon)\right)} = \lim_{r\to\infty} q^{r(N-n)} \mu_{\mathbb{A}^N}\left(\bigcup_{x\in V} B_N(x, q^{-r})\right), \tag{3.1}$$

This limit exists (see [105, 147]) and the map $\mu_X$ thus defined[2] is a measure on the Borel $\sigma$-algebra of $X$.

**Remark 3.2.** When $X \subset \mathbb{P}^{N-1}$ is a projective manifold, one can still define a volume measure $\mu_X$ in the same manner by replacing the measure $\mu_{\mathbb{A}^N}$ in (3.1) with its normalized push-forward to the projective space $\mathbb{P}^{N-1}$, and by defining balls in $\mathbb{P}^{N-1}$ using *Fubini–Study metric*. Here we focus on the affine case and delay our treatment of projective manifolds until Section 3.4.

Given a function $f : X \to \mathbb{R}$ that is integrable with respect to the measure $\mu_X$, we wish to:

1. Estimate the integral $\int_X f(x)\mu_X(dx)$.

[1] The usual notion of smoothness and irreducibility from algebraic geometry.

[2] We can equivalently define a volume measure on $X$ using local charts and differential forms in the usual way.

2. Sample a random variable $\boldsymbol{\xi} \in X$ with the probability density $f(x)/\int_X f(x)\mu_X(dx)$, when $f$ is non-negative and $\int_X f(x)\mu_X(dx) > 0$.

Our sampling method entails intersecting the manifold $X$ with affine linear spaces of complementary dimension. For a matrix $A \in K^{n\times N}$ and $b \in K^n$, we denote by $\mathcal{L}_{A,b}$ the affine linear space implicitly defined as follows:

$$\mathcal{L}_{A,b} \coloneqq \{x \in \mathbb{A}^N \colon Ax = b\}.$$

Such a space is generically of dimension $N-n$ because the set of $(A,b) \in K^{n\times N} \times K^n$ for which $\mathcal{L}_{A,b}$ has dimension $N-n$, is non-empty and Zariski-open. The intersection $\mathcal{L}_{A,b} \cap X$, where $A \in K^{n\times N}$ and $b \in K^n$, is generically finite and its size is bounded by the degree of $X$ (see Section 3.2.1). This is because the set of points $(A,b) \in K^{n\times N} \times K^n$ for which the intersection is finite, is non-empty and Zariski-open. Loosely speaking, sampling from $X$ is then reduced to sampling a random plane $\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}}$ and then sampling a random point from the finite intersection $\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}} \cap X$. However, given a target probability density on $X$, neither sampling step is entirely straightforward. For that reason, we must introduce a *weight* function $w_X \colon X \to \mathbb{R}_{>0}$ on $X$. Before we do so however, we need to define two quantities it involves.

**Definition 3.3.** Let $a, b \geq 1$ be two positive integers, and $M \in K^{a\times b}$ a matrix. As in Proposition 1.24, let us write the Smith normal form of $M$ as $M = UDV$, where $U \in \mathrm{GL}(a,\mathcal{O})$; $V \in \mathrm{GL}(b,\mathcal{O})$; and $D = \mathrm{diag}(\varpi^{v_1},\dots,\varpi^{v_{\min(a,b)}}) \in K^{a\times b}$ with $v_1 \geq \cdots \geq v_{\min(a,b)} \in \mathbb{Z} \cup \{\infty\}$. We then define the *absolute determinant* of $M$ as follows:

$$\mathbf{N}(M) \coloneqq |\varpi^{v_1} \cdots \varpi^{\min(a,b)}| = q^{-v_1-\cdots-v_{\min(a,b)}}.$$

If $E, F$ are $K$-vector spaces of respective dimensions $b, a$ and $\varphi \colon E \to F$ is $K$-linear, we define

$$\mathbf{N}(\varphi) = \mathbf{N}(A),$$

where $A \in K^{a\times b}$ is a matrix representing $\varphi$ in orthonormal bases of $E$ and $F$ in the sense of Section 1.1.3.

**Definition 3.4.** Let $X \subset \mathbb{A}^N$ be an affine algebraic manifold of dimension $n$, and $x$ be a point on $X$. Let $U \in \mathrm{GL}(N,\mathcal{O})$ such that $Ux = (0,\dots,\varpi^{\mathrm{val}(x)})^\top$, and let $W \in \mathcal{O}^{N\times n}$ be a matrix whose columns form an orthonormal basis of the tangent space $T_xX$. Finally, let us set $S_x = \mathrm{diag}(1,\dots,1,\varpi^{\max(0,-\mathrm{val}(x))}) \in K^{N\times N}$. Then we define

$$\mathbf{Nr}(X,x) \coloneqq \mathbf{N}(S_xUW).$$

It is not so clear that this definition does not depend on the choice of $W$ and $U$, but we shall see in Lemma 3.15 that this is the case. The quantity $\mathbf{Nr}(X,x)$ can be interpreted as a "measure" of how far the point $x$ is from being an $\mathcal{O}$-point of $X$.

Now we are ready to define a weight function $w_X$ on $X$.

**Definition 3.5.** Let $X \subset \mathbb{A}^N$ be an affine algebraic manifold of dimension $n$ over $K$. For a point $x \in X$, we define the *weight* of $x$ in $X$ as follows

$$w_X(x) = \frac{1 - q^{-(n+1)}}{1 - q^{-1}} \frac{\max(1, \|x\|^n)}{\mathbf{Nr}(X, x)}.$$

**Remark 3.6.**

1. Notice that, on the $\mathcal{O}$-points $X \cap \mathcal{O}^N$ of the manifold $X$, we have $\mathbf{Nr}(X, x) = 1$ so the weight function $w_X$ is constant and takes the value

$$w_X(x) = \frac{1 - q^{-(n+1)}}{1 - q^{-1}}, \quad x \in X \cap \mathcal{O}^N.$$

2. The weight function $w_X$ is not intrinsic. As we shall see in Proposition 3.16, it depends on how the random linear space $\boldsymbol{\mathcal{L}}_{\boldsymbol{A},\boldsymbol{b}}$ is distributed, or more precisely on the distribution of the random element $(\boldsymbol{A}, \boldsymbol{b}) \in K^{n \times N} \times K^n$.

Given a real valued function $f$ on $X$, we define the following function on $K^{n \times N} \times K^n$:

$$\overline{f}(A, b) \coloneqq \sum_{x \in X \cap \mathcal{L}_{A,b}} w_X(x) f(x), \qquad \text{for } (A, b) \in K^{n \times N} \times K^n,$$

where the sum is 0 by convention when $X \cap \mathcal{L}_{A,b}$ is empty or infinite. Our first result deals with integrating a real-valued integrable function $f$ on a manifold $X$. Namely, we show that the integral can be expressed as the expectation of a real-valued random variable that we can sample. As a matter of notation, random objects shall be bold-faced throughout this chapter.

**Theorem 3.7.** *Let $X \subset \mathbb{A}^N$ be an $n$-dimensional affine algebraic manifold defined over $K$. Let $(\boldsymbol{A}, \boldsymbol{b})$ be a random variable in $K^{n \times N} \times K^n$ with distribution $1_{A \in \mathcal{O}^{n \times N},\ b \in \mathcal{O}^n} dA db$. Then we have:*

$$\int_X f(x) \mu_X(dx) = \mathbb{E}\left[\overline{f}(\boldsymbol{A}, \boldsymbol{b})\right].$$

With this theorem in hand, one can evaluate integrals, up to a certain confidence interval, using Monte-Carlo methods. We discuss this in more detail in Section 3.5.

Our second result deals with sampling a random point $\boldsymbol{\xi}$ from a manifold $X$ with a prescribed probability density $f$ with respect to the natural volume measure $\mu_X$ on $X$:

**Theorem 3.8.** *Let $X \subset \mathbb{A}^N$ be an $n$-dimensional affine algebraic manifold defined over $K$. Let $f \colon X \to \mathbb{R}_{\geq 0}$ be a probability density with respect to $\mu_X$. Let $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{b}})$ be the random variable in $K^{n \times N} \times K^n$ with distribution*

$$\overline{f}(A, b)\ 1_{A \in \mathcal{O}^{n \times N},\ b \in \mathcal{O}^n}\ dA db.$$

*Let $\boldsymbol{\xi}$ be the random variable obtained by intersecting $X$ with the random space $\mathcal{L}_{\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}}}$ and choosing a point $x$ in the finite set $X \cap \mathcal{L}_{\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}}}$ with probability*

$$\frac{w_X(x)f(x)}{\overline{f}(\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}})}.$$

*Then $\boldsymbol{\xi}$ has density $f$ with respect to the volume measure $\mu_X$ on $X$.*

We give similar results for projective manifolds, namely Theorems 3.22 and 3.23, in Section 3.4. We provide an implementation (in SageMath [142]) of the sampling method we describe in this chapter (in particular cases) in following repository:

https://mathrepo.mis.mpg.de/SamplingpAdicManifolds/index.html (3.2)

**Remark 3.9.** Although most results in this chapter stated for algebraic manifolds, there is no issue working with an irreducible variety $X$ (affine or projective) with potential singularities. This is because the singular locus $X^{\text{sing}}$ is lower dimensional in $X$ and we can work with the algebraic manifold $X \setminus X^{\text{sing}}$. Our sampling method will then produce a point in $X$ that is smooth with probability 1.

This chapter is organized as follows. In Section 3.2, we collect some necessary background on $p$-adic varieties and preliminaries on differential geometry. In Section 3.3 we prove the main results, i.e. Theorems 3.7 and 3.8. In Section 3.4 is we discuss the case of projective varities. Finally, in Section 3.5 we discuss how to sample in practice and go over some examples and applications in Section 3.6.

## 3.2 Preliminaries

### 3.2.1 A pinch of intersection theory

In this section we recall some facts from intersection theory of algebraic varieties (see [86, Section 18] and [47] for more details). The reader may skip this and come back to it when necessary.

Let $X \subset \mathbb{A}^N$ be an affine algebraic manifold of dimension $n$ and let $d$ be its degree. Then there exists a variety $\mathcal{V}_X$ in $K^{n\times N} \times K^n$ of lower dimension such that

$$\#(\mathcal{L}_{A,b} \cap X) \leq d \quad \text{for any } (A,b) \in (K^{n\times N} \times K^n) \setminus \mathcal{V}_X. \tag{3.3}$$

Since $\mathcal{V}_X$ is a lower dimensional variety in $K^{n\times N} \times K^n$ it has measure 0 with respect to the volume measure $dAdb$. So if $(\boldsymbol{A},\boldsymbol{b})$ is a random variable that has a density with respect to $dAdb$, then with probability 1, the intersection $\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}} \cap X$ contains at most $d$ points. So, given a real valued function $f\colon X \to \mathbb{R}$ the function

$$\overline{f}\colon (A,b) \mapsto \sum_{x \in \mathcal{L}_{A,b}\cap X} w_X(x)f(x),$$

is well defined on the Zariski open set $K^{n\times N} \times K^n \setminus \mathcal{V}_X$.

Similarly, if $X \subset \mathbb{P}^{N-1}$ is a projective algebraic $K$-manifold of dimension $n$ and degree $d$, then the set of matrices $A \in K^{n\times N}$ such that the intersection $\mathcal{L}_A \cap X$ is infinite, where

$$\mathcal{L}_A = \{x \in \mathbb{P}^{N-1} \colon Ax = 0\},$$

is a lower dimensional algebraic variety $\mathcal{W}_X$ in $K^{n\times N}$. Hence $\mathcal{W}_X$ has measure 0 with respect to the Haar measure $dA$ and

$$\#(\mathcal{L}_A \cap X) \leq d \quad \text{for any } A \in K^{n\times N} \setminus \mathcal{W}_X. \tag{3.4}$$

Moreover, given a real valued function $f \colon X \to \mathbb{R}$ the function

$$\overline{f} \colon A \mapsto (1+q^{-1}) \cdots (1+q^{-n}) \sum_{x \in \mathcal{L}_A \cap X} f(x)$$

is well defined on the Zariski open set $K^{n\times N} \setminus \mathcal{W}_X$.

### 3.2.2 The $p$-adic co-area formula

In this section we recall a few notions on $p$-adic integration on manifolds. We refer the reader to [26, 105, 145, 134] and references therein for a more detailed account.

Let $X$ be a smooth algebraic (affine or projective) manifold defined over $K$. One can then endow the variety $X$ with the structure of a $K$-analytic manifold in the sense of [26] and a volume measure $\mu_X$. A definition of the latter is given in Equation (3.1) for the affine case and in Equation (3.8) for the projective case.

**Definition 3.10.** Let $X$ and $Y$ be two $K$-analytic manifolds, $x \in X$ and $\varphi : X \to Y$ be a $K$-analytic map. We define the *absolute Jacobian* of $\varphi$ at $x$ as

$$\mathbf{J}(\varphi, x) \coloneqq \mathbf{N}(\mathbf{D}_x \varphi),$$

the absolute determinant of the $K$-linear map $\mathbf{D}_x \varphi \colon T_x X \to T_{\varphi(x)} Y$.

The following is the $p$-adic coarea formula from [26].

**Theorem 3.11** ([26, Theorem 6.2.1])**.** *Let $X$ and $Y$ be two analytic $K$-manifolds with $\dim(X) \geq \dim(Y)$ and let $\varphi : X \to Y$ be a $K$-analytic map. Then, for any function $f \colon X \to \mathbb{R}$ that is integrable with respect to the volume measure on $X$, we have*

$$\int_X \mathbf{J}(\varphi, x) f(x) \mu_X(dx) = \int_Y \left( \int_{\varphi^{-1}(y)} f(z) \mu_{\varphi^{-1}(y)}(dz) \right) \mu_Y(dy).$$

**Corollary 3.12.** *Let $X$ and $Y$ be two $K$-analytic manifolds and $\varphi\colon X \to Y$ an analytic map from $X$ to $Y$.*

*(i) Suppose that $\boldsymbol{\xi}$ is an $X$-valued random variable with density $f$ with respect to $\mu_X$. Then the density $g$ of $\boldsymbol{\eta} = \varphi(\boldsymbol{\xi})$ with respect to $\mu_Y$ is*

$$g(y) = \int_{\varphi^{-1}(y)} \frac{f(z)}{\mathbf{J}(\varphi, z)} \mu_{\varphi^{-1}(y)}(dz).$$

*(ii) Let $\boldsymbol{\eta}$ be a $Y$-valued random variable with density $g$ with respect to $\mu_Y$ and let $\boldsymbol{\xi}$ be the $X$-valued random variable such that, conditioned on $(Y = y)$, the variable $\boldsymbol{\xi}$ has density $f_y$ on $\varphi^{-1}(Y)$ with respect to $\mu_{\varphi^{-1}(y)}$. Then $\boldsymbol{\xi}$ has density*

$$f(x) = \mathbf{J}(\varphi, x) g(\varphi(x)) f_{\varphi(x)}(x).$$

*Proof.* (i) Let $V$ be a Borel set in $Y$. Applying Theorem 3.11 we get

$$\begin{aligned} P(\boldsymbol{\eta} \in V) &= \int_X 1_V(\varphi(x)) f(x) \mu_X(dx) \\ &= \int_Y \left( \int_{\varphi^{-1}(y)} \frac{f(z)}{\mathbf{J}(\varphi, z)} 1_V(\varphi(z)) \mu_{\varphi^{-1}(y)}(dz) \right) \mu_Y(dy) \\ &= \int_Y \left( \int_{\varphi^{-1}(y)} \frac{f(z)}{\mathbf{J}(\varphi, z)} \mu_{\varphi^{-1}(y)}(dz) \right) 1_V(y) \mu_Y(dy) \\ &= \int_Y g(y) \mu_Y(dy). \end{aligned}$$

(ii) Let $U$ be a Borel set in $X$. Then, applying Theorem 3.11 we get

$$\begin{aligned} P(\boldsymbol{\xi} \in U) &= \mathbb{E}[P(\boldsymbol{\xi} \in U | \boldsymbol{\eta})] \\ &= \int_Y \left( \int_{\varphi^{-1}(y)} f_y(z) 1_U(z) \mu_{\varphi^{-1}(y)}(dz) \right) g(y) \mu_Y(dy) \\ &= \int_Y \left( \int_{\varphi^{-1}(y)} g(\varphi(z)) f_{\varphi(z)}(z) 1_U(z) \mu_{\varphi^{-1}(y)}(dz) \right) \mu_Y(dy) \\ &= \int_X \mathbf{J}(\varphi, x) g(\varphi(x)) f_{\varphi(x)}(x) 1_U(x) \mu_X(dx) \\ &= \int_X f(x) 1_U(x) \mu_X(dx). \end{aligned}$$

□

We denote by $\mathrm{Gr}(n, K^m)$ the Grassmannian variety parametrizing $n$-dimensional vector subspaces of $K^m$. The orthogonal group $\mathrm{GL}(m, \mathcal{O})$ has a natural action on $\mathrm{Gr}(n, K^m)$.

**Lemma 3.13.** *Let $m \geq n \geq 1$ be two integers. There exists is a unique orthogonally invariant probability distribution on the Grassmanian $\mathrm{Gr}(n, K^m)$.*

*Proof.* Since $\mathrm{GL}(m, \mathcal{O})$ acts transitively on $\mathrm{Gr}(n, K^m)$ and the stabilizer of the subspace generated by the first $n$ vectors of the standard basis of $K^m$ is

$$H = \left\{ \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right) : A \in \mathrm{GL}(k, \mathcal{O}), B \in \mathrm{GL}(m-n, \mathcal{O}) \quad \text{and} \quad C \in \mathcal{O}^{n\times(m-n)} \right\},$$

we can write $\mathrm{Gr}(n, K^m)$ as a homogeneous space as follows

$$\mathrm{Gr}(n, K^m) = \mathrm{GL}(m, \mathcal{O})/H.$$

Let $\nu$ be a probability measure on $\mathrm{GL}(m, \mathcal{O})/H$ that is $\mathrm{GL}(m, \mathcal{O})$-invariant. Then its pull-back $\nu^*$ to $\mathrm{GL}(m, \mathcal{O})$ is also $\mathrm{GL}(m, \mathcal{O})$ invariant, so it is a Haar measure on $\mathrm{GL}(m, \mathcal{O})$ with $\nu^*(\mathrm{GL}(m, \mathcal{O})) = 1$. We then conclude since $\mathrm{GL}(m, \mathcal{O})$ is a compact topological group, there is a unique Haar measure on $\mathrm{GL}(m, \mathcal{O})$ up to scaling. □

We end this section with the following simple but useful lemma.

**Lemma 3.14.** *Let $n \geq 1$ be a positive integer, then we have*

$$\int_{C\in\mathcal{O}^{n\times n}} |\det(C)| dC = \frac{1-q^{-1}}{1-q^{-(n+1)}}.$$

*Proof.* Let $\boldsymbol{C}$ be a random matrix in $K^{n\times n}$ whose entries are independent and uniform in $\mathcal{O}$. Then the integral in question is the expectation $\mathbb{E}[|\det(\boldsymbol{C})|]$. We can compute this expectation using the distribution of $|\det(\boldsymbol{C})|$ from [62, Theorem 4.1]. We then have

$$\mathbb{E}[|\det(\boldsymbol{C})|] = (1-q^{-1})\cdots(1-q^{-n}) \sum_{m=0}^{\infty} \binom{n+m-1}{m}_{q^{-1}} q^{-2m},$$

where $\binom{n}{k}_{q^{-1}}$ denotes the usual $q^{-1}$-binomial coefficient (also known as the Gaussian binomial coefficient):

$$\binom{n}{k}_{q^{-1}} := \frac{(1-q^{-1})\cdots(1-q^{-n})}{(1-q^{-1})\cdots(1-q^{-k}) \times (1-q^{-1})\cdots(1-q^{-(n-k)})} \quad \text{for } n \geq k \geq 0$$

Then using the well known generating series

$$\sum_{m=0}^{\infty} \binom{n+m-1}{m}_{q^{-1}} t^m = \prod_{k=0}^{n-1} \frac{1}{1-q^{-k}t},$$

we get

$$\sum_{m=0}^{\infty} \binom{n+m-1}{m}_{q^{-1}} q^{-2m} = \prod_{k=0}^{n-1} \frac{1}{1-q^{-k-2}} = \frac{1}{(1-q^{-2})\cdots(1-q^{-(n+1)})}.$$

So we deduce that

$$\mathbb{E}[|\det(\boldsymbol{C})|] = \int_{C\in\mathcal{O}^{n\times n}} |\det(C)| dC = \frac{1-q^{-1}}{1-q^{-(n+1)}}.$$ □

## 3.3 Sampling from affine manifolds

In this section, we proceed to proving the main results of this chapter, namely Theorems 3.7 and 3.8. Similar results for projective manifolds are stated and proved in Section 3.4. We start with the following:

**Lemma 3.15.** *Let $X \subset \mathbb{A}^N$ be an affine algebraic $K$-manifold of dimension $n$. The definition of $\mathbf{Nr}(X, x)$ in Definition 3.4 does not depend on the choice of $U$ and $W$.*

*Proof.* Set $S_x = \mathrm{diag}(1, \dots, 1, \varpi^{\max(0, -\operatorname{val}(x))})$ and let $U_1, U_2 \in \mathrm{GL}(N, \mathcal{O})$ and $W_1, W_2 \in \mathcal{O}^{N \times n}$ be such that

(1) $U_1 x = U_2 x = (0, \dots, 0, \varpi^{\operatorname{val}(x)})^\top$,

(2) columns of $W_1, W_2$ are two orthonormal bases of the tangent space $T_x X$.

Then there exists $V \in \mathrm{GL}(n, \mathcal{O})$ such that $W_2 = W_1 V$. Let $A = U_2 U_1^{-1}$ and $B = S_x A S_x^{-1}$ so that we have $B(S_x U_1 W_1)V = S_x U_2 W_2$. We claim that $B \in \mathrm{GL}(N, \mathcal{O})$ is an orthogonal matrix. To see why, notice that, thanks to condition (1), the matrix $A$ is of the form

$$A = \left(\begin{array}{ccc|c} & & & 0 \\ & A' & & \vdots \\ & & & 0 \\ \hline z_1 & \cdots & z_{N-1} & 1 \end{array}\right)$$

where $A' \in \mathrm{GL}(N-1, \mathcal{O})$ and $z_1, \dots, z_{N-1} \in \mathcal{O}$. We then deduce that $B$ is of the form

$$B = \left(\begin{array}{ccc|c} & & & 0 \\ & A' & & \vdots \\ & & & 0 \\ \hline \alpha z_1 & \cdots & \alpha z_{N-1} & 1 \end{array}\right)$$

where $\alpha = \varpi^{\max(0, -\operatorname{val}(x))} \in \mathcal{O}$. So we deduce that $B \in \mathrm{GL}(N, \mathcal{O})$. Now, since $V$ and $B$ are both orthogonal, from Definition 3.3 we can see that

$$\mathbf{N}(S_x U_1 W_1) = \mathbf{N}(S_x U_2 W_2) = \mathbf{N}(B S_x U_1 W_1 V),$$

which finishes the proof. □

This means that the weight function $w_X$ in Definition 3.5 is indeed well defined.

**Proposition 3.16.** *Let $X \subset \mathbb{A}^N$ be an affine algebraic $K$-manifold of dimension $n$ and $x$ a point on $X$. We then have*

$$w_X(x) = \left( \int_{A \in K^{n \times N}} |\det(A_{|T_x X})| 1_{A \in \mathcal{O}^{n \times N}, \|Ax\| \leq 1} dA \right)^{-1}.$$

*Proof.* Let $U \in \mathrm{GL}(N, \mathcal{O})$ such that $y := Ux = (0, \dots, 0, \varpi^{\mathrm{val}(x)})^\top$. Let $W$ be a matrix whose columns form an orthonormal basis of $T_x X$. Let us fix the matrix $R_x = \mathrm{diag}(1, \dots, 1, \varpi^{v(x)}) \in K^{N\times N}$. Let us denote by $I_X(x)$ the following integral

$$I_X(x) = \int_{A\in K^{n\times N}} |\det(A_{|T_x X})| 1_{A\in\mathcal{O}^{n\times N}} 1_{Ax\in\mathcal{O}^n} dA.$$

Then, by a change of variable $BU = A$, we have

$$\begin{aligned} I_X(x) &= \int_{B\in K^{n\times N}} |\det((BU)_{|T_x X})| 1_{B\in\mathcal{O}^{n\times N}} 1_{By\in\mathcal{O}^n} dB \\ &= \int_{B\in K^{n\times N}} |\det(BUW)| 1_{B\in\mathcal{O}^{n\times N}} 1_{BR_x\in\mathcal{O}^{n\times N}} dB \\ &= \int_{B\in\mathcal{O}^{n\times N}\cap(\mathcal{O}^{n\times N}R_x^{-1})} |\det(BUW)| dB. \end{aligned}$$

Notice the following equality

$$\mathcal{O}^{n\times N} \cap \mathcal{O}^{n\times N} R_x^{-1} = \mathcal{O}^{n\times N} S_x,$$

where $S_x = \mathrm{diag}(1, \dots, 1, \varpi^{\max(0, -\mathrm{val}(x))}) \in K^{N\times N}$. So, using the change of variables $B = B'S_x$, we deduce that

$$\begin{aligned} I_X(x) &= \int_{B\in\mathcal{O}^{n\times N}S_x} |\det(BUW)| dB \\ &= \left(\frac{1}{\max(1, \|x\|)}\right)^n \int_{B'\in\mathcal{O}^{n\times N}} |\det(B'S_x UW)| dB'. \end{aligned}$$

Let us write the Smith normal form of the matrix $S_x UW$, i.e.

$$S_x UW = V_1 D V_2,$$

where $V_1 \in \mathrm{GL}(N, \mathcal{O}), V_2 \in \mathrm{GL}(n, \mathcal{O})$ and $D = \mathrm{diag}(\varpi^{v_1}, \dots, \varpi^{v_n}) \in K^{N\times n}$. So, by the change of variables $B'V_1 = C$, we get

$$\begin{aligned} I_X(x) &= \frac{1}{\max(1, \|x\|^n)} \int_{B'\in\mathcal{O}^{n\times N}} |\det(B'V_1 D V_2)| dB' \\ &= \frac{1}{\max(1, \|x\|^n)} \int_{C\in\mathcal{O}^{n\times N}} |\det(CD)| dC \\ &= \frac{q^{-(v_1+\cdots+v_n)}}{\max(1, \|x\|^n)} \int_{C\in\mathcal{O}^{n\times n}} |\det(C)| dC. \end{aligned}$$

Combining the previous equation with Definition 3.4, Definition 3.5 and Lemma 3.14, we get

$$I_X(x) = \frac{\mathbf{Nr}(X, x)}{\max(1, \|x\|^n)} \frac{1 - q^{-1}}{1 - q^{-(n+1)}} = \frac{1}{w_X(x)}$$

as desired. □

**Remark 3.17.** (i) Proposition 3.16 gives another proof of the fact that $\mathbf{Nr}(X, x)$ does not depend on the choice of $U$ and $W$ in Definition 3.4.

(ii) Recall, from Remark 3.6, that the weight function is constant on $X \cap \mathcal{O}^N$. Unwinding the definition of $w_X$ we can also see that for $U \in \mathrm{GL}(N, \mathcal{O})$ and $x \in X$ we have $w_{UX}(Ux) = w_X(x)$ where $UX = \{Ux \colon x \in X\}$.

(iii) If the probability density $f$ we wish to sample from is supported on $X \cap \varpi^{-r}\mathcal{O}^N$, we can scale $X$ by $\varpi^r$ and sample $\boldsymbol{\xi}'$ from $\varpi^r X \cap \mathcal{O}^N$ (where the weight function is constant) with density $f(\varpi^{-r}\cdot)$. We can then obtain a random variable $\boldsymbol{\xi}$ on $X$ with density $f$ by taking $\boldsymbol{\xi} = \varpi^{-r}\boldsymbol{\xi}'$.

We are now ready to prove our main theorems.

*Proof of Theorem 3.7.* By definition we have

$$\mathbb{E}(\overline{f}(\boldsymbol{A}, \boldsymbol{b})) = \int_{K^{n\times N}\times K^n} \overline{f}(A, b) 1_{A\in\mathcal{O}^{n\times N}, b\in\mathcal{O}^n} dA db.$$

Let us define the following map:

$$\varphi \colon K^{n\times N} \times X \to K^{n\times N} \times K^n, \quad (A, x) \mapsto (A, Ax). \tag{3.5}$$

The map $\varphi$ is analytic and its differential is given by

$$D_{(A,x)}\varphi \colon K^{n\times N} \times T_x X \to K^{n\times N} \times K^n, \quad (H, u) \mapsto (H, Hx + Au).$$

So the differential of $\varphi$ at $(A, x)$ acts trivially on the first component of the product $K^{n\times N} \times T_x X$ and acts as $A_{|T_x X}$ on the second, i.e. its determinant is given by

$$\mathbf{J}(\varphi, (A, x)) = \left|\det(A_{|T_x X})\right|. \tag{3.6}$$

Applying Theorem 3.11 for the function $\varphi$ yields

$$\begin{aligned}
&\int_{K^{n\times N}\times X} \left|\det(A_{|T_x X})\right| w_X(x) f(x) 1_{A\in\mathcal{O}^{n\times N},\ Ax\in\mathcal{O}^n}\ dA\ \mu_X(dx) \\
&= \int_{K^{n\times N}\times X} \mathbf{J}(\varphi, (A, x)) w_X(x) f(x) 1_{A\in\mathcal{O}^{n\times N},\ Ax\in\mathcal{O}^n}\ dA\ \mu_X(dx) \\
&= \int_{K^{n\times N}\times K^n} \left( \int_{\varphi^{-1}(A,y)} w_X(z) f(z) \mu_{\varphi^{-1}(A,y)}(dz) 1_{A\in\mathcal{O}^{n\times N},\ Az\in\mathcal{O}^n}\ \mu_{\varphi^{-1}(A,y)(dz)} \right) dA dy \\
&= \int_{K^{n\times N}\times K^n} \left( \int_{\varphi^{-1}(A,y)} w_X(z) f(z) \mu_{\varphi^{-1}(A,y)}(dz) \right) 1_{A\in\mathcal{O}^{n\times N},\ y\in\mathcal{O}^n}\ dA dy.
\end{aligned}$$

But, for $A \in K^{n\times N}$ and $y \in K^n$ we have

$$\varphi^{-1}((A, y)) = \{(A, z) \in K^{n\times N} \times X \colon Az = y \text{ and } z \in X\},$$

and this is a finite set for almost every $A$ and $y$. So for almost every $A$ and $y$, the measure $\mu_{\varphi^{-1}(A,y)}$ equals the counting measure on the finite set $\varphi^{-1}((A,y))$ and we then have

$$\begin{aligned}
&\int_{K^{n\times N}\times X} |\det(A_{|T_xX})| w_X(x) f(x) 1_{A\in\mathcal{O}^{n\times N},\ Ax\in\mathcal{O}^n}\ dA\mu_X(dx)\\
=&\int_{K^{n\times N}\times K^n} \left(\sum_{\substack{x\in X,\\ Ax=y}} w_X(x)f(x)\right) 1_{A\in\mathcal{O}^{n\times N}, y\in\mathcal{O}^n}\ dAdy\\
=&\int_{K^{n\times N}\times K^n} \overline{f}(A,y) 1_{A\in\mathcal{O}^{n\times N},\ y\in\mathcal{O}^n}\ dAdy\\
=&\mathbb{E}[\overline{f}(\boldsymbol{A},\boldsymbol{b})].
\end{aligned}$$

Hence the equation

$$\mathbb{E}\left[\overline{f}(\boldsymbol{A},\boldsymbol{b})\right] = \int_X \left(\int_{K^{n\times N}} |\det(A_{|T_xX})| 1_{A\in\mathcal{O}^{n\times N},\ Ax\in\mathcal{O}^n} dA\right) w_X(x) f(x) \mu_X(dx). \tag{3.7}$$

Then, combining Equation (3.7) and Proposition 3.16, we conclude that

$$\mathbb{E}\left[\overline{f}(\boldsymbol{A},\boldsymbol{b})\right] = \int_X f(x)\mu_X(dx). \qquad \square$$

*Proof of Theorem 3.8.* Let $(\widetilde{\boldsymbol{A}},\boldsymbol{\xi})$ be the random variable, with values in $K^{n\times N}\times X$, obtained by first sampling $(\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}}) \in K^{n\times N}\times K^n$ from with distribution $\overline{f}(A,b)1_{A\in\mathcal{O}^{n\times N},b\in\mathcal{O}^n}dAdb$ and then choosing a point $\boldsymbol{\xi}$ from $\mathcal{L}_{\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}}}\cap X$ with probability

$$\frac{w_X(x)f(x)}{\overline{f}(\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}})}.$$

Then applying Corollary 3.12-(ii) to the map $\varphi$ from Equation (3.5), we deduce that $(\widetilde{\boldsymbol{A}},\boldsymbol{\xi})$ has density

$$\begin{aligned}
g_{(\widetilde{\boldsymbol{A}},\boldsymbol{\xi})}(A,x) &= \overline{f}\left(\varphi(A,x)\right) 1_{\varphi(A,x)\in\mathcal{O}^{n\times N}\times\mathcal{O}^n} \frac{w_X(x)f(x)}{\overline{f}(\varphi(A,x))} \mathbf{J}(\varphi,(A,x))\\
&= w_X(x)f(x) 1_{\varphi(A,x)\in\mathcal{O}^{n\times N}\times\mathcal{O}^n} \mathbf{J}(\varphi,(A,x))
\end{aligned}$$

with respect to the volume measure $dA\mu_X(dx)$ on $K^{n\times N}\times X$. Computing the second marginal of this joint distribution, we deduce that the density $g_{\boldsymbol{\xi}}$ of $\boldsymbol{\xi}$ is

$$\begin{aligned}
g_{\boldsymbol{\xi}}(x) &= \int_{A\in K^{n\times N}} w_X(x)f(x)\,\mathbf{J}(\varphi,(A,x)) 1_{\varphi(A,x)\in\mathcal{O}^{n\times N}\times\mathcal{O}^n} dA\\
&= w_X(x)f(x) \int_{A\in K^{n\times N}} |\det(A_{|T_xX})| 1_{A\in\mathcal{O}^{n\times N}, Ax\in\mathcal{O}^n} dA\\
&= f(x).
\end{aligned}$$

The second (resp. third) equation follows from Equation (3.6) (resp. Proposition 3.16). So, as desired, $\boldsymbol{\xi}$ has density $f$ with respect to $\mu_X$ on $X$. $\square$

## 3.4 Sampling from projective manifolds

This section deals with sampling from projective algebraic manifolds. More precisely, we shall state and prove analogs of Theorem 3.7 and Theorem 3.8 in projective space.

Let $N \geq 2$ be an integer. We denote by $\mathbb{P}^{N-1}$ the projective space of dimension $N-1$ over $K$. Let us denote by $S^{N-1}$ the unit sphere in $K^N$, i.e.,

$$S^{N-1} \coloneqq \{x \in K^N \colon \|x\| = 1\}.$$

We warn the reader that, unlike the Euclidean setting, the unit sphere is actually an open set in $K^N$ and has dimension $N$ (as a topological space). Consider the *Hopf fibration*

$$\psi \colon S^{N-1} \to \mathbb{P}^{N-1}, \quad (x_1, \ldots, x_N) \mapsto (x_1 \colon \cdots \colon x_N).$$

The projective space $\mathbb{P}^{N-1}$ can be endowed with a metric $d$ defined as follows:

$$d(x, y) = \|\tilde{x} \wedge \tilde{y}\|, \quad x, y \in \mathbb{P}^{N-1}$$

where $\tilde{x} \in \psi^{-1}(x)$, $\tilde{y} \in \psi^{-1}(y)$ and the norm $\|\tilde{x} \wedge \tilde{y}\|$ is the standard norm in $\bigwedge^2 K^N$ associated to its standard lattice $\bigwedge^2 \mathcal{O}^N$. This metric is called the *Fubini-Study* metric. For $x \in \mathbb{P}^{N-1}$ and $\epsilon > 0$ let us denote by

$$\mathbb{B}_{N-1}(x, \epsilon) \coloneqq \{y \in \mathbb{P}^{N-1} \colon d(x, y) \leq \epsilon\}$$

the ball of radius $\epsilon$ around $x$.

Endowed with the metric $d$, the projective space $\mathbb{P}^{N-1}$ is a compact metric space on which we define a volume measure $\mu_{\mathbb{P}^{N-1}}$ as follows

$$\mu_{\mathbb{P}^{N-1}} \coloneqq \frac{1}{1 - q^{-1}} \psi_* \mu_{S^{N-1}},$$

that is the normalized push-forward of $\mu_{S^{N-1}}$ by $\psi$[3]. The measure $\mu_{\mathbb{P}^{N-1}}$ is finite and we have

$$\mu_{\mathbb{P}^{N-1}}(\mathbb{P}^{N-1}) = \frac{1 - q^N}{1 - q^{-1}}.$$

**Remark 3.18.** Notice that from Lemma 3.14, we have

$$\int_{C \in \mathcal{O}^{n \times n}} |\det(C)| dC = \frac{1}{\mu_{\mathbb{P}^n}(\mathbb{P}^n)}.$$

[3]Notice that $\mu_{S^{N-1}}(S^{N-1}) = 1 - q^{-N}$.

A projective algebraic variety in $\mathbb{P}^{N-1}$ is the zero set of a system of homogeneous polynomials $\boldsymbol{p} = (p_1, \ldots, p_r)$ in $K[x_1, \ldots, x_N]$; that is

$$\{x \in \mathbb{P}^{N-1} \colon p_1(x) = \cdots = p_r(x) = 0\}.$$

We refer to irreducible and smooth projective varieties as projective algebraic manifolds.

Let $X \subset \mathbb{P}^{N-1}$ be an algebraic projective manifold of dimension $n \geq 1$. Similar to the affine case (3.1), we can define a volume measure on $X$ as follows:

$$\mu_X(V) \coloneqq \lim_{\epsilon \to 0} \frac{\mu_{\mathbb{P}^{N-1}}\left(\bigcup_{x \in V} \mathbb{B}_{N-1}(x, \epsilon)\right)}{\mu_{\mathbb{P}^{N-1-n}}\left(\mathbb{B}_{N-1-n}(0, \epsilon)\right)}, \qquad \text{for } V \subset X \text{ open.} \tag{3.8}$$

The limit in (3.8) exists (see [147, 105] for more details) and this defines a volume measure $\mu_X$ on the projective manifold $X$.

**Remark 3.19.** For our purposes, the main difference between the affine and projective spaces is that the projective space is a compact topological space (with the quotient topology induced by the Hopf fibration $\psi$). So, unlike the affine case, a projective algebraic manifold admits a uniform probability density. Also, loosely speaking, there are no "far" points in the projective space, so as we shall see, the weight function is constant or, in other words, no point gets more weight than another. We can say that the space is, in some sense, "isotropic".

Before we state our results for projective manifolds, we recall a few facts and establish a couple of preliminary results.

### 3.4.1 Preliminaries

Suppose that $X \subset \mathbb{P}^{N-1}$ is a projective algebraic manifold of dimension $n$ defined by homogeneous polynomials $p_1, \ldots, p_r \in K[x_1, \ldots, x_N]$ and let $x$ be a point in $X$. The tangent space $T_x X$ can be defined in many ways, and one way to do so is the following. The cone $\widetilde{X} \subset \mathbb{A}^N$ over $X$ defined as follows

$$\widetilde{X} = \{(\lambda y_1, \ldots, \lambda y_N) \in \mathbb{A}^N \colon \lambda \in K \text{ and } (y_1 : \cdots : y_N) \in X\}.$$

This is an affine algebraic variety which is smooth at every non-zero point $x \in \widetilde{X} \setminus \{0\}$ and has dimension $n+1$. The tangent space $T_x\widetilde{X}$ is a linear subspace in $K^N$ of dimension $n+1$ and $x \in T_x\widetilde{X}$. The tangent space $T_x X$ can then be defined as an orthogonal complement[4] of the line $K \cdot x$ in $T_x\widetilde{X}$ and we thus view $T_x X$ as a linear subspace[5] of $K^N$ of dimension $n$.

[4] All such vector spaces are isomorphic to one another.

[5] The projective tangent space is also often defined as the projectivisation of $T_x\widetilde{X}$.

**Proposition 3.20.** *Let $X \subset \mathbb{P}^{N-1}$ be a projective algebraic manifold of dimension $n$ and let us define $\mathcal{X} \subset \mathbb{A}^{n\times N} \times \mathbb{P}^{N-1}$ as follows:*

$$\mathcal{X} = \{(A, x) \in \mathbb{A}^{n\times N} \times X \colon Ax = 0\}.$$

*Then $\mathcal{X}$ is a manifold, and for $(A, x) \in \mathcal{X}$ we have*

$$T_{(A,x)}\mathcal{X} = \{(H, h) \in K^{n\times N} \times T_xX \colon Hx + Ah = 0\}.$$

*Moreover, if $\varphi, \phi$ are the projections from $\mathcal{X}$ to $\mathbb{A}^{n\times N}$ and $\mathbb{P}^{N-1}$ respectively, then we have*

$$\frac{\mathbf{J}(\varphi, (A, x))}{\mathbf{J}(\phi, (A, x))} = |\det(A_{|T_xX})|,$$

*for $(A, x) \in \mathcal{X}$ such that $A_{|T_xX}$ is an isomorphism.*

*Proof.* Let $(p_1, \dots, p_r) \in K[x_1, \dots, x_N]$ be homogeneous polynomials generating the ideal of $X$. Let $(A, x) \in \mathcal{X}$ and let $J_x$ be the following Jacobian matrix

$$J_x = \left(\frac{\partial p_i}{\partial x_j}(x)\right)_{1\le i\le r, 1\le j\le N}.$$

Then, considering $\mathcal{X}$ as the variety in $\mathbb{A}^{n\times N} \times \mathbb{P}^{N-1}$ cut out by the equations $Ax = 0$ and $p_1(x) = \cdots = p_r(x) = 0$ we can compute the Jacobian matrix of $\mathcal{X}$ at the point $(A, x)$. This matrix represents the linear map

$$K^{n\times N} \times K^N \to K^n \times K^r, \quad (H, h) \mapsto (Hx + Ah, J_xh).$$

The tangent space of $\mathcal{X}$ at $(A, x)$ is the kernel of this map, so

$$T_{(A,x)}\mathcal{X} = \{(H, h) \in K^{n\times N} \times T_xX \colon Hx + Ah = 0\}.$$

The projection maps $\varphi, \phi$ are clearly analytic, and for any $(A, x) \in \mathcal{X}$ we have

$$\begin{aligned} d_{(A,x)}\varphi \colon T_{(A,x)}\mathcal{X} &\to K^{n\times N} & \qquad d_{(A,x)}\phi \colon T_{(A,x)}\mathcal{X} &\to T_xX \\ (H, h) &\mapsto H & (H, h) &\mapsto h. \end{aligned}$$

Suppose that $(A, x) \in \mathcal{X}$ is such that $A_{|T_xX}$ is an isomorphism. Fix $U \in \mathrm{GL}(N, \mathcal{O})$ such that $Ux = (1 : \cdots : 0)^\top$ and define the maps

$$\begin{aligned} \pi_1 \colon K^{n\times N} &\to T_{(A,x)}\mathcal{X} & \qquad \pi_2 \colon T_xX &\to T_{(A,x)}\mathcal{X} \\ H &\mapsto (H, -(A_{|T_xX})^{-1}Hx) & h &\mapsto ((-Ah|0)U, h) \end{aligned}$$

where $(-Ah|0) \in K^{n\times N}$. Notice that $d_{(A,x)}\varphi \circ \pi_1 = \mathrm{Id}_{K^{n\times N}}$ and $d_{(A,x)}\phi \circ \pi_2 = \mathrm{Id}_{T_xX}$. Since $A \in \mathcal{O}^{n\times N}$ we have

$$\mathbf{N}(\pi_2) = 1,$$

because $\pi_2$ sends any orthonormal basis of $T_xX$ to an orthonormal family in $T_{(A,x)}\mathcal{X} \subset K^{n\times N} \times K^N$. Also, since $A \in \mathcal{O}^{N\times N}$, the singular values of $A_{|T_xX}$ are all in $\mathcal{O}$ so the singular values of $A^{-1}_{|T_xX}$ have negative or zero valuation. From this we can see that

$$\mathbf{N}(\pi_1) = |\det(A_{|T_xX})|^{-1}.$$

We deduce that

$$\frac{\mathbf{J}(\varphi,(A,x))}{\mathbf{J}(\phi,(A,x))} = \frac{\mathbf{N}(\pi_2)}{\mathbf{N}(\pi_1)} = |\det(A_{|T_xX})|. \qquad \square$$

**Lemma 3.21.** *Let $X$ be a projective manifold of dimension $n$ in $\mathbb{P}^{N-1}$ and $x$ be a point on $X$. Set $M_x := \{A \in K^{n\times N} : Ax = 0\}$. Then*

$$\int_{M_x} |\det(A_{|T_xX})| 1_{A\in\mathcal{O}^{n\times N}} \mu_{M_x}(dA) = \frac{1-q^{-1}}{1-q^{-(n+1)}}.$$

*Proof.* Let $W \in \mathcal{O}^{N\times n}$ be a matrix whose columns form an orthonormal basis of $T_xX$ and let $U \in \mathrm{GL}(N,\mathcal{O})$ such that $Ux = e_1 = (1:0:\cdots:0)^\top$. The space $M_x$ is a vector space of dimension $(N-1)\times n$ and $M_{e_1}U = M_x$. So, with the change of variable $BU = A$, we get

$$\begin{aligned}\int_{M_x} |\det(A_{|T_xX})| 1_{A\in\mathcal{O}^{n\times N}} \mu_{M_x}(dA) &= \int_{A\in M_x} |\det(AW)| 1_{A\in\mathcal{O}^{n\times N}} \mu_{M_x}(dA)\\ &= \int_{B\in M_{e_1}} |\det(BUW)| 1_{B\in\mathcal{O}^{N\times n}}\ \mu_{M_x}(dA)\\ &= \int_{C\in\mathcal{O}^{n\times(N-1)}} |\det((0\mid C)UW)| dC.\end{aligned}$$

Let $\widetilde{W} \in K^{(N-1)\times n}$ be the matrix obtained from $UW$ by deleting the first row and let us write the Smith normal form of $\widetilde{W}$ as

$$\widetilde{W} = V_1 D V_2,$$

where $V_1 \in \mathrm{GL}(N-1,\mathcal{O}), V_2 \in \mathrm{GL}(n,\mathcal{O})$ and $D = \mathrm{diag}(\varpi^{v_1},\dots,\varpi^{v_n}) \in K^{(N-1)\times n}$. We then deduce that

$$\begin{aligned}\int_{M_x} |\det(A_{|T_xX})| 1_{A\in\mathcal{O}^{n\times N}} \mu_{M_x}(dA) &= \int_{C\in\mathcal{O}^{n\times(N-1)}} |\det(C\widetilde{W})| dC\\ &= q^{-(v_1+\cdots+v_n)} \int_{C\in\mathcal{O}^{n\times N}} |\det(C)| dC\\ &= \mathbf{N}(\widetilde{W}) \frac{1-q^{-1}}{1-q^{-(n+1)}}.\end{aligned}$$

Since $X$ is a projective manifold, the tangent space $T_xX$ is orthogonal to $x$ (see Section 3.4.1). We deduce that the columns of $UW$ are orthogonal to $(1,\dots,0)^\top$ so $\widetilde{W}$ has orthonormal columns. Hence $\mathbf{N}(\widetilde{W}) = 1$ which finishes the proof. $\square$

Similarly to the affine case given a real valued function $f\colon X \to \mathbb{R}$, we define the weighted average function of $f$ as follows:

$$\overline{f}(A) = \sum_{x \in \mathcal{L}_A \cap X} f(x), \quad \text{for } A \in K^{n\times N}.$$

By convention, the sum is taken to be 0 whenever $\mathcal{L}_A \cap X$ is empty or infinite.

### 3.4.2 Sampling from projective manifolds

Now we state and prove the analogues of Theorems 3.7 and 3.8 for the projective case.

**Theorem 3.22.** *Let $X \subset \mathbb{P}^{N-1}$ be an $n$-dimensional projective algebraic manifold defined over $K$. Let $\boldsymbol{A}$ be a random variable in $K^{n\times N}$ with distribution $1_{A\in\mathcal{O}^{n\times N}}dA$. Then we have*

$$\int_X f(x)\mu_X(dx) = \mu_{\mathbb{P}^n}(\mathbb{P}^n)\ \mathbb{E}[\overline{f}(\boldsymbol{A})],$$

*with*

$$\mu_{\mathbb{P}^n}(\mathbb{P}^n) = \frac{1-q^{-(n+1)}}{1-q^{-1}}$$

*Proof.* Let $\mathcal{X} \subset \mathbb{A}^{n\times N} \times \mathbb{P}^{N-1}$ be the algebraic variety defined by

$$\mathcal{X} \coloneqq \left\{(A,x) \in \mathbb{A}^{n\times N} \times X\colon Ax = 0\right\}.$$

Let us denote by $\varphi$ and $\phi$ the natural projections from $\mathcal{X}$ onto $K^{n\times N}$ and $X$ respectively, and, for a point $x \in X$, set $M_x \coloneqq \{A \in K^{n\times N}\colon Ax = 0\}$. We apply Theorem 3.11 on $\varphi$ and then on $\phi$ to get the following

$$\begin{aligned}
\mathbb{E}\left[\overline{f}(\boldsymbol{A})\right] &= \int_{K^{n\times N}} \left(\sum_{\substack{x\in X,\\ Ax=0}} f(x)\right) 1_{A\in\mathcal{O}^{n\times N}} dA \\
&= \int_{K^{n\times N}} \left( \int_{(A,z)\in\varphi^{-1}(A)} f(z) 1_{A\in\mathcal{O}^{n\times N}} \mu_{\varphi^{-1}(A)}(dz)\right) dA \\
&= \int_{\mathcal{X}} \mathbf{J}(\varphi,(A,x)) f(x) 1_{A\in\mathcal{O}^{n\times N}} \mu_{\mathcal{X}}(dA,dx) \\
&= \int_X \left( \int_{(A,x)\in\phi^{-1}(x)} \frac{\mathbf{J}(\varphi,(A,x))}{\mathbf{J}(\phi,(A,x))} 1_{A\in\mathcal{O}^{n\times N}} \mu_{\phi^{-1}(dA)}\right) f(x)\mu_X(dx) \\
&= \int_X \left( \int_{A\in M_x} |\det(A_{|T_xX})| 1_{A\in\mathcal{O}^{n\times N}} \mu_{\phi^{-1}(x)}(dA)\right) f(x)\mu_X(dx) \\
&= \frac{1-q^{-1}}{1-q^{-(n+1)}} \int_X f(x)\mu_X(dx) \\
&= \frac{1}{\mu_{\mathbb{P}^n}(\mathbb{P}^n)} \int_X f(x)\mu_X(dx).
\end{aligned}$$

The last equality follows from Lemma 3.21. □

**Theorem 3.23.** *Let $X \subset \mathbb{P}^N$ be an $n$-dimensional projective algebraic manifold defined over $K$. Let $f\colon X \to \mathbb{R}_{\geq 0}$ be a probability density with respect to the volume measure $\mu_X$ on $X$. Let $\widetilde{\boldsymbol{A}}$ be the random variable in $K^{n\times N}$ with distribution*

$$\frac{1-q^{-(n+1)}}{1-q^{-1}}\overline{f}(A)1_{A\in\mathcal{O}^{n\times N}}dA.$$

*Let $\boldsymbol{\xi}$ be the random variable obtained by intersecting $X$ with the random space $\mathcal{L}_{\widetilde{\boldsymbol{A}}}$ and choosing a point $x$ in the finite set $X\cap\mathcal{L}_{\widetilde{\boldsymbol{A}}}$ with probability*

$$\frac{f(x)}{\overline{f}(\widetilde{\boldsymbol{A}})}.$$

*Then $\boldsymbol{\xi}$ has density $f$ with respect to $\mu_X$.*

*Proof.* Let $(\widetilde{\boldsymbol{A}},\boldsymbol{\xi})$ be the random variable with values in $\mathcal{X}$ (as defined in Proposition 3.20) such that $\widetilde{\boldsymbol{A}}$ has distribution

$$\overline{f}(A)1_{A\in\mathcal{O}^{n\times N}}dA$$

and, given $\widetilde{\boldsymbol{A}}$, $\boldsymbol{\xi}$ is a random point in $\mathcal{L}_{\widetilde{\boldsymbol{A}}}\cap X$ with probability

$$P(\boldsymbol{\xi}=x|\widetilde{\boldsymbol{A}})=\frac{f(x)}{\overline{f}(\widetilde{\boldsymbol{A}})}.$$

Then, by virtue of Corollary 3.12-(ii) applied to the projection map $\varphi:\mathcal{X}\to K^{n\times N}$, we deduce that the density of $(\widetilde{\boldsymbol{A}},\boldsymbol{\xi})$, with respect to $\mu_{\mathcal{X}}$, is given by

$$\begin{aligned}f_{\widetilde{\boldsymbol{A}},\boldsymbol{\xi}}(A,x)&=\frac{1-q^{-(n+1)}}{1-q^{-1}}\overline{f}(A)1_{A\in\mathcal{O}^{n\times N}}\frac{f(x)}{\overline{f}(A)}\,\mathbf{J}(\varphi,(A,x))\\&=\frac{1-q^{-(n+1)}}{1-q^{-1}}f(x)\,\mathbf{J}(\varphi,(A,x)),\end{aligned}$$

for $(A,x)\in\mathcal{X}$. Applying Corollary 3.12-(i) of to the projection map $\phi:\mathcal{X}\to X$, we then deduce that the density of $\boldsymbol{\xi}$ is

$$\begin{aligned}f_{\boldsymbol{\xi}}(x)&=\int_{\phi^{-1}(x)}\frac{1-q^{-(n+1)}}{1-q^{-1}}f(x)\frac{\mathbf{J}(\varphi,(A,x))}{\mathbf{J}(\phi,(A,x))}\mu_{\phi^{-1}(x)}(dA)\\&=\frac{1-q^{-(n+1)}}{1-q^{-1}}f(x)\int_{\phi^{-1}(x)}\frac{\mathbf{J}(\varphi,(A,x))}{\mathbf{J}(\phi,(A,x))}\mu_{\phi^{-1}(x)}(dA)\\&=\frac{1-q^{-(n+1)}}{1-q^{-1}}f(x)\int_{\phi^{-1}(x)}|\det(A_{|T_xX})|\mu_{\phi^{-1}(x)}(dA)\\&=f(x).\end{aligned}$$

The last equation follows from the Lemma 3.21. This concludes the proof. □

## 3.5 Sampling linear spaces in practice

In this section we explain how to sample the random planes $\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}}$ and $\mathcal{L}_{\boldsymbol{A}}$ explicitly. We also explain how to sample the random planes $\mathcal{L}_{\widetilde{\boldsymbol{A}},\widetilde{\boldsymbol{b}}}$ from Theorem 3.8 and $\mathcal{L}_{\widetilde{\boldsymbol{A}}}$ from Theorem 3.23 by rejection sampling, and we give bounds on how efficient this sampling method is.

### 3.5.1 Sampling linear spaces explicitly

When the codimension of the manifold $X$ is small (hypersurfaces for example), for computational reasons, it is easier to find the intersection of $X$ with a linear space $\mathcal{E}$ of complementary dimension $N-n$ when the latter has an explicit form. That is writing $\mathcal{E}$ in the form

$$\mathcal{E} = u + \operatorname{span}_K(x_1, \dots, x_{N-n}),$$

where $u \in K^N$ and $x_1, \dots, x_{N-n} \in K^N$ are linearly independent.

**Lemma 3.24.** *Let $\boldsymbol{A} \in K^{n\times N}$, $\boldsymbol{b} \in K^n$ and $\boldsymbol{B} \in K^{(N+1)\times(N-n+1)}$ be matrices with random i.i.d entries uniformly distributed in $\mathcal{O}$, and $\boldsymbol{u}, \boldsymbol{x_1}, \dots, \boldsymbol{x_{N-n}} \in K^N$ be such that*

$$\begin{pmatrix}\boldsymbol{u}\\1\end{pmatrix}, \begin{pmatrix}\boldsymbol{x_1}\\0\end{pmatrix}, \dots, \begin{pmatrix}\boldsymbol{x_{N-n}}\\0\end{pmatrix} \text{ form an orthonormal basis of } \operatorname{columnspan}(\boldsymbol{B}).$$

*The random affine space $\mathcal{E}_{\boldsymbol{u},\boldsymbol{x_1},\dots,\boldsymbol{x_{N-n}}} := \boldsymbol{u} + \operatorname{span}(\boldsymbol{x_1}, \dots, \boldsymbol{x_{N-n}})$ has the same probability distribution as $\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}}$.*

*Proof.* Notice that the linear space $\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}}$ can be written as

$$\mathcal{L}_{\boldsymbol{A},\boldsymbol{b}} = \left\{ x \in K^N \colon (\boldsymbol{A}| - \boldsymbol{b}) \begin{pmatrix}x\\1\end{pmatrix} = 0 \right\}.$$

So it suffices to show that $\operatorname{columnspan}(\boldsymbol{B})$ and $\operatorname{Ker}((\boldsymbol{A}| - \boldsymbol{b}))$ have the same distribution in the Grassmannian $\operatorname{Gr}(N-n+1, K^{N+1})$. Thanks to Lemma 3.13, it is enough to notice that the distributions of $\operatorname{columnspan}(\boldsymbol{B})$ and $\operatorname{Ker}((\boldsymbol{A}|-\boldsymbol{b}))$ are both orthogonally invariant. This is indeed the case since for any $U \in \operatorname{GL}(N+1, \mathcal{O})$ we have[6]

$$(\boldsymbol{A}| - \boldsymbol{b})U \overset{d}{=} (\boldsymbol{A}| - \boldsymbol{b}) \quad \text{and} \quad U\boldsymbol{B} \overset{d}{=} \boldsymbol{B}. \qquad \square$$

### 3.5.2 Rejection sampling

Let $X \subset \mathbb{A}^N$ be an affine algebraic manifold of dimension $n$ and degree $d$ and let $f \colon X \to \mathbb{R}_{\geq 0}$ be a probability density function with respect to $\mu_X$. We recall that the average function

[6] By "$\overset{d}{=}$" we mean equality in distribution.

$\overline{f}$ in the affine case is defined as

$$\overline{f}(A,b) = \sum_{x \in \mathcal{L}_{A,b} \cap X} w_X(x) f(x), \qquad \text{for } (A,b) \in K^{n\times N} \times K^n,$$

where, by convention, the sum is 0 whenever the intersection $\mathcal{L}_{A,b} \cap X$ is empty or infinite.

**Proposition 3.25** (Rejection sampling)**.** *Suppose that there exists a constant $M > 0$ such that $\overline{f}(A,b) < M$ almost everywhere with respect to $dAdb$. Let $(\boldsymbol{A}, \boldsymbol{b})$ be the random variable with distribution $1_{A\in\mathcal{O}^{n\times N}, b\in\mathcal{O}^n} dAdb$ and let $\boldsymbol{\eta}$ be a random variable such that*

$$P(\boldsymbol{\eta} = 1 | (\boldsymbol{A}, \boldsymbol{b})) = \frac{\overline{f}(\boldsymbol{A}, \boldsymbol{b})}{M} \quad \text{and} \quad P(\boldsymbol{\eta} = 0 | (\boldsymbol{A}, \boldsymbol{b})) = \frac{M - \overline{f}(\boldsymbol{A}, \boldsymbol{b})}{M}.$$

*Then, conditioned on the event $(\boldsymbol{\eta} = 1)$, the random variable $(\boldsymbol{A}, \boldsymbol{b})$ has distribution*

$$\overline{f}(A,b) 1_{A\in\mathcal{O}^{n\times N}, b\in\mathcal{O}^n} dAdb.$$

*Proof.* This follows directly from Bayes' rule as follows

$$\begin{aligned}
P\big((\boldsymbol{A}, \boldsymbol{b}) \in (dA, db) | \eta = 1\big) &= \frac{P\big(\eta = 1 | (\boldsymbol{A}, \boldsymbol{b}) \in (dA, db)\big) P\big((\boldsymbol{A}, \boldsymbol{b}) \in (dA, db)\big)}{P(\boldsymbol{\eta} = 1)} \\
&= \frac{P\big(\eta = 1 | (\boldsymbol{A}, \boldsymbol{b}) \in (dA, db)\big)}{P(\boldsymbol{\eta} = 1)} 1_{A\in\mathcal{O}^{n\times N},\ b\in\mathcal{O}^n}\ dAdb \\
&= \frac{\overline{f}(A,b)/M}{\mathbb{E}[\overline{f}(\boldsymbol{A}, \boldsymbol{b})]/M} 1_{A\in\mathcal{O}^{n\times N},\ b\in\mathcal{O}^n}\ dAdb \\
&= \overline{f}(A,b).
\end{aligned}$$

The last equation follows from Theorem 3.7 and the equality

$$P(\boldsymbol{\eta} = 1) = \mathbb{E}[P(\boldsymbol{\eta} = 1 | (\boldsymbol{A}, \boldsymbol{b}))] = \frac{1}{M} \mathbb{E}[\overline{f}(\boldsymbol{A}, \boldsymbol{b})] = \frac{1}{M}.$$ □

**Lemma 3.26.** *Let $f\colon X \to \mathbb{R}_{\geq 0}$ be a probability density function supported on $X \cap \varpi^{-r}\mathcal{O}^N$ for some integer $r \geq 0$. Suppose that $\kappa \coloneqq \sup_{x\in X} f(x) < \infty$. Then we have*

$$\overline{f}(A,b) \leq d q^{(n+1)r}\ \frac{1 - q^{-(n+1)}}{1 - q^{-1}} \kappa.$$

*In particular, if $f$ is the uniform probability density on $X \cap \mathcal{O}^N$ then*

$$\overline{f}(A,b) \leq d\ \frac{1 - q^{-(n+1)}}{1 - q^{-1}}.$$

*Proof.* Let $x \in X$ and $U, W, S_x$ as in Definition 3.4. Then, since the columns of $UW$ are orthonormal in $K^N$, its rows are in $\mathcal{O}^n$ and, modulo $\varpi$, they span $k^n$. So we deduce that

$$\mathbf{Nr}(X, x) = \mathbf{N}(S_x UW) \geq \min(1, \|x\|^{-1}).$$

Hence, from Definition 3.5, we get

$$w_X(X, x) \leq \frac{1 - q^{-(n+1)}}{1 - q^{-1}} \max(1, \|x\|^{n+1}).$$

Then for $(A, b) \in K^{n \times N} \times K^n$ we get

$$\overline{f}(A, b) \leq \#(X \cap \mathcal{L}_{A,b}) \frac{1 - q^{-(n+1)}}{1 - q^{-1}} q^{(n+1)r} \sup_{x \in X} f(x).$$

Since the number of intersection points $\#(X \cap \mathcal{L}_{A,b})$ is at most $d = \deg(X)$ (except for a measure zero set of $(A, b) \in K^{n \times N} \times K^n$, see Section 3.2.1), we deduce the desired result. The second statement is an immediate consequence of the first one. □

**Remark 3.27.** The bound given for $\overline{f}(A, b)$ is far from being sharp. Moreover, when one wishes to sample from $X \cap \varpi^{-r}\mathcal{O}^N$, this bound is not very practical for rejection sampling. In this case, it is better to use Remark 3.17 (iii).

Let $h \colon X \to \mathbb{R}$ be an integrable function on $X$ supported on $X \cap \mathcal{O}^N$ and let $(\boldsymbol{A_i}, \boldsymbol{b_i})_{i \geq 0}$ be a sequence of i.i.d random variables such that $(\boldsymbol{A_i}, \boldsymbol{b_i})$ has the uniform distribution on $\mathcal{O}^{n \times N} \times \mathcal{O}^n$ for all $i \geq 0$. Finally, set

$$S_m(h) \coloneqq \overline{h}(\boldsymbol{A_1}, \boldsymbol{b_1}) + \overline{h}(\boldsymbol{A_2}, \boldsymbol{b_2}) + \cdots + \overline{h}(\boldsymbol{A_m}, \boldsymbol{b_m}).$$

Then we have the following:

**Proposition 3.28.** *The random variable $S_m(h)/m$ converges almost surely to the integral $I(h) \coloneqq \int_X h(x)\mu_X(dx)$ as $m \uparrow \infty$. Moreover, if $\kappa \coloneqq \sup_{x \in X} |h(x)| < \infty$, then*

$$P\left(\left|\frac{S_m(h)}{m} - I(h)\right| \geq \epsilon\right) \leq \frac{\kappa^2 d^2}{\epsilon^2 m} \left(\frac{1 - q^{-(n+1)}}{1 - q^{-1}}\right)^2, \quad \text{for } m \geq 1.$$

*Proof.* The first statement is an immediate application of the law of large numbers. The second follows from Lemma 3.26 and Chebychev's inequality. □

**Remark 3.29.** While this section focuses on affine manifolds, the results discussed within can be restated and proved for projective manifolds without much difficulty.

## 3.6 Applications and examples

In this section we discuss a few concrete examples and applications. The first case of interest is when the algebraic manifold $X$ is an algebraic group.

### 3.6.1 Measures on algebraic groups

Let $G$ be an algebraic group defined over $K$, by which we mean a smooth (either affine or projective) algebraic variety together with

1. (*identity element*) an element $e \in G$,
2. (*multiplication*) a morphism $m\colon G \times G \to G, (x, y) \mapsto xy$,
3. (*inverse*) a morphism $\iota\colon G \to G, x \mapsto x^{-1}$,

with respect to which $G$ is a group (see [123] or [19] for a detailed account). In our discussion, $m$ and $\iota$ are $K$-morphisms and we are interested in the group $G(K)$ of $K$ points of $G$ which we also denote by $G$ for simplicity and, for our purposes, $G$ is embedded in some affine or projective space over $K$.

The group $G$ is a locally compact topological group and thus admits a left Haar measure; that is a non-zero measure $\nu_G$ such that

$$\nu(gA) = \nu(A), \quad \text{for any Borel measurable set } A \subset G.$$

which is unique up to scaling. If $G$ is an algebraic group embedded in a projective space as an algebraic manifold, then $G$ is compact and the measure $\mu_G$ is then finite and also right-invariant. In this case we normalize $\nu$ so that $\nu(G) = 1$. In the case where $G$ is affine, the measure $\nu$ is finite on the set $G(\mathcal{O})$ of $\mathcal{O}$-points of $G$ and we normalise $\nu$ so that $\nu(G(\mathcal{O})) = 1$.

**Remark 3.30.** It is not always the case that the points in $G(\mathcal{O})$ form a subgroup of $G$. For example, this fails to be the case for $G = \mathrm{GL}(n, K)$.

**Example 3.31.** Let $n \geq 1$ be a positive integer. If $G$ is either the special linear group $\mathrm{SL}(n, K)$ or the special orthogonal group $\mathrm{SO}(n, K)$ or the symplectic group $\mathrm{Sp}(n, K)$, the $\mathcal{O}$-points $G(\mathcal{O})$ form a compact subgroup of $G$. Moreover, the normalized Haar measure $\nu$ on $G(\mathcal{O})$ coincides with the uniform probability measure on $G(\mathcal{O})$ with respect to the volume measure $\mu_G$ (as defined in Equation (3.1)). This is because the measure $\mu_{\mathbb{A}^{n\times n}}$ is invariant under the action of $\mathrm{GL}(n, \mathcal{O})$ and in particular under the action of $G(\mathcal{O}) \subset \mathrm{GL}(n, \mathcal{O})$ and hence $\mu_G$ is also $G(\mathcal{O})$-invariant. So using Theorem 3.8 we can sample from the Haar measure on the compact matrix groups $\mathrm{SL}(n, \mathcal{O})$, $\mathrm{SO}(n, \mathcal{O})$ and $\mathrm{Sp}(n, \mathcal{O})$. For small values of $n$, we provide examples of this in the repository (3.2).

In general however, the measure $\mu_G$ (as defined in Equation (3.1) or Equation (3.8)) may not be invariant under the action of $G$. In other words, the following may fail:

$$\mu_G(g \cdot A) = \mu_G(A), \quad \text{for any Borel set } A \subset G.$$

Note that the measure $\mu_G$ depends on how $G$ is embedded in its ambient space.

### 3.6.2 Moduli spaces

Another case of interest is when the algebraic manifold $X$ is a moduli space parametrizing certain objects. Then sampling from $X$, we can get an idea of how often a certain property of these objects holds or how rare are objects of certain kind are in $X$. We give two examples of such a situation.

#### 3.6.2.1 Modular curves

Let $N$ be a positive integer and consider the modular curve $X_1(N)$. This is a smooth projective curve defined over $\mathbb{Q}$, and it has the following *moduli* interpretation: for any field $K$ with characteristic 0, noncuspidal[7] $K$-points of $X_1(N)$ classify isomorphism classes of pairs $(E, P)$, where $E$ is an elliptic curve over $K$ and $P$ is a point of $E(K)$ of order $N$. For the theory of modular curves, see [42]. See also [152, Section C.13] for a quick introduction.

In this example, we will sample uniformly from $\mathbb{Z}_{31}$-points of $X_1(30)$, and compute the *Tamagawa numbers* of the corresponding elliptic curves over $\mathbb{Q}_{31}$. For an elliptic curve $E/\mathbb{Q}_p$, the finite index

$$c_p = [E(\mathbb{Q}_p) : E^0(\mathbb{Q}_p)]$$

is referred to as the Tamagawa number of $E/\mathbb{Q}_p$, where $E^0(\mathbb{Q}_p)$ is the subgroup of $E(\mathbb{Q}_p)$ consisting of points that have good reduction. Clearly, if $E/\mathbb{Q}_p$ has good reduction, then $c_p$ equals 1. We note that Tamagawa numbers of elliptic curves are important local arithmetic invariants. They arise in the conjecture of Birch and Swinnerton-Dyer, for example; see [152, Section C.16]. Moreover, they can be easily computed using Magma [20].

The following (optimized) equation for $X_1(30)$ was provided by Sutherland in [158]:

$$\begin{aligned} X_1(30) : y^6 &+ (x^6 - 5x^5 + 6x^4 + 3x^3 - 6x^2 + 7x + 3)y^5 \\ &+ (x^7 - 3x^6 - 13x^5 + 44x^4 - 18x^3 + x^2 + 18x + 3)y^4 \\ &+ (x^8 - 3x^7 - 13x^6 + 27x^5 + 46x^4 - 32x^3 + 21x^2 + 15x + 1)y^3 \\ &+ 2x(x^7 - 8x^6 + 9x^5 + 20x^4 + 6x^3 - 6x^2 + 9x + 2)y^2 \\ &- 4x^2(2x^5 - 7x^4 - 3x^3 - 1)y + 8x^6 = 0. \end{aligned}$$

Moreover, if $(x_0, y_0)$ is a noncuspidal point on $X_1(30)$, then the corresponding elliptic curve is of the form

$$y^2 = x^3 + (t^2 - 2qt - 2)x^2 - (t^2 - 1)(qt + 1)^2 x,$$

where

$$\begin{aligned} q &= y_0 + 1, \\ t &= 4(y_0 + 1)(x_0 + y_0)/(x_0 y_0^3 - 4x_0 y_0 - 4x_0 - 3y_0^3 - 6y_0^2 - 4y_0). \end{aligned}$$

[7] Modular curves have only finitely many cuspidal points. This will be important for what follows.

See the table in https://math.mit.edu/~drew/X1_optcurves.html. Table 3.1 presents the Tamagawa numbers of elliptic curves obtained for a sample of 500.000 $\mathbb{Z}_{31}$-points on $X_1(30)$, and the number of times they occurred.

| $c_{31}$ | | $c_{31}$ | | $c_{31}$ | |
|---|---|---|---|---|---|
| 1 | 266775 | 8 | 1 | 20 | 382 |
| 2 | 53317 | 9 | 48 | 24 | 2 |
| 3 | 56726 | 10 | 13174 | 30 | 6549 |
| 4 | 1601 | 12 | 1804 | 45 | 16 |
| 5 | 27759 | 15 | 12956 | 60 | 192 |
| 6 | 58623 | 18 | 68 | 90 | 7 |

Table 3.1: The Tamagawa numbers and their multiplicities that appeared in our sampling.

#### 3.6.2.2 Hilbert modular surfaces

Here, we will work with *Hilbert modular surfaces* $Y_-(D)$, with the notation in [56]. These surfaces parametrize abelian surfaces with real multiplication. More precisely, let $d > 1$ be a squarefree integer, and set

$$D = \begin{cases} d & \text{if } d \equiv 1 \bmod 4, \\ 4d & \text{if } d \equiv 2, 3 \bmod 4. \end{cases}$$

Note that $D$ is nothing but the discriminant of the ring of integers $\mathcal{O}_D$ of the real quadratic field $\mathbb{Q}(\sqrt{D})$. Such a number is called a *positive fundamental discriminant*. The quotient

$$\mathrm{PSL}_2(\mathcal{O}_D) \setminus \left(\mathcal{H}^+ \times \mathcal{H}^-\right)$$

is the coarse moduli space of principally polarized abelian surfaces with real multiplication by $\mathcal{O}_D$. Here, $\mathcal{H}^+$ (resp. $\mathcal{H}^-$) denotes the complex upper (resp. lower) half plane. There is a holomorphic map from this quotient to the moduli space $\mathcal{A}_2$ of principally polarized abelian surfaces. The image is the *Humbert surface* $\mathcal{H}_D$, and the Hilbert modular surface $Y_-(D)$ is a double cover of $\mathcal{H}_D$ branched along a finite union of modular curves. For the theory of Hilbert modular surfaces, see, for example, [77, 24].

The surfaces $Y_-(D)$ have models over $\mathbb{Q}$, and points on these surfaces correspond, generically, to Jacobians of smooth projective curves[8] of genus 2. Explicit equations for birational models of $Y_-(D)$, as well as the Igusa–Clebsch invariants $I_2, I_4, I_6$ and $I_{10}$ of the corresponding genus-2 curves, were provided by Elkies and Kumar in [56] for all fundamental discriminants $D$ between 1 and 100. In this final example, we will

[8]Recall that a principally polarized abelian surface over an algebrically closed field is either the Jacobian variety of a smooth projective curve of genus 2 or the product of two elliptic curves.

- sample uniformly from $\mathbb{Z}_5$-points of $Y_-(5)$, and
- compute the minimal skeleta of the Berkovich analytifications of the corresponding genus-2 curves.

It is well know that there are precisely 7 different (graph-theoretical) types, which are depicted in Figure 3.2. The recent work of Helminck [89] shows that *tropical Igusa invariants*, which can easily be computed from Igusa–Clebsch invariants, distinguish between the different types; see [89, Theorem 2.11]. See also Chapter 7 for a similar result concerning Picard curves.

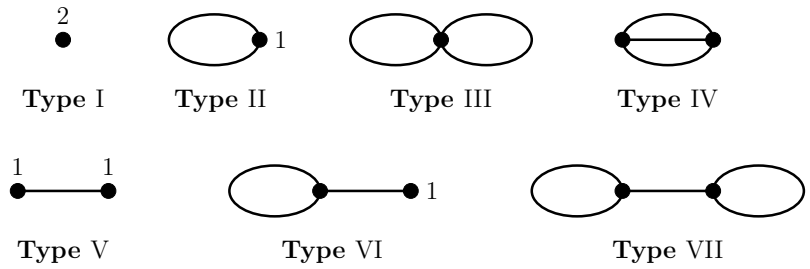



Figure 3.2: Minimal skeleta of the Berkovich analytifications of genus-2 curves.

A birational model of the surface $Y_-(5)$ is given by

$$z^2 = 2\left(6250h^2 - 4500g^2h - 1350gh - 108h - 972g^5 - 324g^4 - 27g^3\right),$$

see [56, Theorem 16]. Moreover, the map from from $\mathcal{H}_5$ to $\mathcal{A}_2$ (or, more precisely, to the moduli space $\mathcal{M}_2$ of curves of genus 2) is given by

$$(I_2 : I_4 : I_6 : I_{10}) = \left(6(4g+1), 9g^2, 9(4h + 9g^3 + 2g^2), 4h^2\right)$$

see [56, Corollary 15]. Table 3.2 shows how many times the types occurred for a sample of 500000 $\mathbb{Z}_5$-points on $Y_-(5)$.

| **Type** I | **Type** II | **Type** III | **Type** IV | **Type** V | **Type** VI | **Type** VII |
|---|---|---|---|---|---|---|
| 414900 | 0 | 40338 | 23040 | 21688 | 2 | 32 |

Table 3.2: The multiplicities of the types that appeared in our sampling.

As shown the in table, Types II, VI and VII are quite rare. In fact, we never see Type II, and it is unclear to the authors if there is a theoretical reason behind this.

## 3.7 Conclusion

In conclusion, this chapter presents a method to sample from $p$-adic algebraic manifolds which is based on slicing with random linear spaces; a technique that has been previously used for real algebraic manifolds. This makes it possible to use probabilistic methods to estimate certain geometric or number theoretic quantities.

# Chapter 4

# The Bernoulli clock

This chapter is based on joint work [53] with Jim Pitman.

## 4.1 Introduction

The *Bernoulli polynomials* $(B_n(x))_{n\geq 0}$ are a special sequence of univariate polynomials with rational coefficients. They are named after the Swiss mathematician Jakob Bernoulli (1654–1705), who (in his *Ars Conjectandi* published posthumously in Basel 1713) found the sum of $m$th powers of the first $n$ positive integers using the instance $x = 1$ of the *power sum formula*

$$\sum_{k=0}^{n-1}(x+k)^m = \frac{B_{m+1}(x+n) - B_{m+1}(x)}{m+1}, \qquad (n = 1, 2, \ldots, m = 0, 1, 2, \ldots). \tag{4.1}$$

The evaluations $B_m := B_m(0)$ and $B_m(1) = (-1)^m B_m$ are known as the *Bernoulli numbers*, from which the polynomials are recovered as

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_{n-k}\, x^k. \tag{4.2}$$

These polynomials have been well studied, starting from the early work of Faulhaber, Bernoulli, Seki and Euler in the 17th and early 18th centuries. They can be defined in multiple ways. For example, Euler defined the Bernoulli polynomials by their *exponential generating function*

$$B(x, \lambda) := \frac{\lambda e^{\lambda x}}{e^{\lambda} - 1} = \sum_{n=0}^{\infty} \frac{B_n(x)}{n!}\lambda^n \qquad (|\lambda| < 2\pi). \tag{4.3}$$

Beyond evaluating power sums, the Bernoulli numbers and polynomials are useful in other contexts and appear in many areas in mathematics, among which we mention number

theory [10, 13, 121, 4], Lie theory [21, 139, 25, 117], algebraic geometry and topology [90, 124] and probability [111, 112, 93, 92, 157, 132, 17].

The *factorially normalized Bernoulli polynomials* $b_n(x) := B_n(x)/n!$ can also be defined inductively as follows (see [125, §9.5]). Beginning with $b_0(x) = B_0(x) = 1$, for each positive integer $n$, the function $x \mapsto b_n(x)$ is the unique antiderivative of $x \mapsto b_{n-1}(x)$ that integrates to 0 over $[0, 1]$:

$$b_0(x) = 1, \quad \frac{d}{dx} b_n(x) = b_{n-1}(x) \quad \text{and} \quad \int_0^1 b_n(x) = 0 \qquad (n > 0). \tag{4.4}$$

So the first few polynomials $b_n(x)$ are

$$\begin{aligned} b_0(x) &= 1, & b_1(x) &= x - 1/2, \\ b_2(x) &= \frac{1}{2!}(x^2 - x - 1/6), & b_3(x) &= \frac{1}{3!}(x^3 - 3x^2/2 + x/2). \end{aligned}$$

As shown in [125, Theorem 9.7] starting from (4.4), the functions $f(x) = b_n(x)$ with argument $x \in [0, 1)$ are also characterized by the simple form of their *Fourier transform*

$$\widehat{f}(k) := \int_0^1 f(x) e^{-2\pi i k x} dx \qquad (k \in \mathbb{Z}) \tag{4.5}$$

which is given by

$$\begin{aligned} &\widehat{b}_0(k) = 1[k = 0], && \text{for } k \in \mathbb{Z}; \\ &\widehat{b}_n(0) = 0 \quad \text{and} \quad \widehat{b}_n(k) = -\frac{1}{(2\pi i k)^n}, && \text{for } n > 0 \text{ and } k \neq 0, \end{aligned} \tag{4.6}$$

with the notation $1[\cdots]$ equal to 1 if $[\cdots]$ holds and 0 otherwise. It follows from the Fourier expansion of $b_n(x)$:

$$b_n(x) = -\frac{2}{(2\pi)^n} \sum_{k=1}^{\infty} \frac{1}{k^n} \cos\left(2k\pi x - \frac{n\pi}{2}\right)$$

that there exists a constant $C > 0$ such that

$$\sup_{0 \le x \le 1} \left| (2\pi)^n b_n(x) + 2 \cos\left(2\pi x - \frac{n\pi}{2}\right) \right| \le C 2^{-n} \quad \text{for } n \ge 2, \tag{4.7}$$

see [108]. So as $n$ tends to $\infty$ the polynomials $b_n(x)$ looks like shifted cosine functions. Besides (4.3) and (4.4), several other characterizations of the Bernoulli polynomials are described in [107, 35].

This chapter draws attention to an explicit construction of the Bernoulli polynomials by *circular convolution*. For a pair of functions $f = f(u)$ and $g = g(u)$, defined for $u$ in $[0, 1)$ identified with the circle group $\mathbb{T} := \mathbb{R}/\mathbb{Z} = [0, 1)$, with $f$ and $g$ integrable with respect to Lebesgue measure on $\mathbb{T}$, their *circular convolution* $f \circledast g$ is the function

$$(f \circledast g)(u) = \int_{\mathbb{T}} f(v) g(u - v) dv \qquad \text{for } u \in \mathbb{T}. \tag{4.8}$$

Here $u - v$ is evaluated in the circle group $\mathbb{T}$, that is modulo 1, and $dv$ is the shift-invariant Lebesgue measure on $\mathbb{T}$ with total measure 1. Iteration of this operation defines the $n$th convolution power $u \mapsto f^{\circledast n}(u)$ for each positive integer $n$, each integrable $f$, and $u \in \mathbb{T}$.

**Theorem 4.1.** *The factorially normalized Bernoulli polynomials* $b_n(x) = \frac{B_n(x)}{n!}$ *are characterized by:*

*(i)* $b_0(x) = 1$ *and* $b_1(x) = x - 1/2$,

*(ii) for* $n > 0$ *the* $n$*-fold circular convolution of* $b_1(x)$ *with itself is* $(-1)^{n-1} b_n(x)$*; that is*

$$b_n(x) = (-1)^{n-1} b_1^{\circledast n}(x). \tag{4.9}$$

In view of the well known expression of circular convolution as multiplication of Fourier transforms $\widehat{f \circledast g} = \widehat{f}\,\widehat{g}$, Theorem 4.1 follows from the classical Fourier evaluation (4.6) and uniqueness of the Fourier transform. A more elementary proof of Theorem 4.1, without Fourier transforms, is provided in Section 4.2. So the Fourier evaluation (4.6) may be regarded as a corollary of Theorem 4.1. That theorem can also be reformulated as follows:

**Corollary 4.2.** *The following identities hold for circular convolution of factorially normalized Bernoulli polynomials:*

$$\begin{aligned} b_0 \circledast b_0 &= b_0, \\ b_0 \circledast b_n &= 0 \quad (n \geq 1), \\ b_n \circledast b_m &= -b_{n+m} \quad (n, m \geq 1). \end{aligned}$$

In particular, for positive integers $n$ and $m$, this evaluation of $(b_n \circledast b_m)(1)$ yields an identity which appears in [130, p. 31]:

$$(-1)^m \int_0^1 b_n(u) b_m(u) du = \int_0^1 b_n(u) b_m(1-u) du = -b_{n+m}(1). \tag{4.10}$$

Here the first equality is due to the well known *reflection symmetry* of the Bernoulli polynomials

$$(-1)^m b_m(u) = b_m(1-u) \qquad (m \geq 0) \tag{4.11}$$

which is the equality of the coefficients of $\lambda^m$ in the elementary identity of Eulerian generating functions

$$B(u, -\lambda) = \frac{(-\lambda) e^{-\lambda u}}{e^{-\lambda} - 1} = \frac{\lambda e^{\lambda(1-u)}}{e^{\lambda} - 1} = B(1-u, \lambda). \tag{4.12}$$

The rest of this chapter is organized as follows. Section 4.2 gives an elementary proof for Theorem 4.1, and discusses circular convolution of polynomials. In Section 4.3 we highlight the fact that $1 - 2^n b_n(x)$ is the probability density at $x \in (0, 1)$ of the fractional part of a

sum of $n$ independent random variables, each with the beta$(1,2)$ probability density $2(1-x)$ at $x \in (0,1)$. Because the minimum of two independent uniform $[0,1]$ variables has this beta$(1,2)$ probability density, the circular convolution of $n$ independent beta$(1,2)$ variables is closely related to a continuous model we call the *Bernoulli clock*: Spray the circle $\mathbb{T} = [0,1)$ of circumference 1 with $2n$ i.i.d uniform positions $U_1, U_1', \ldots, U_n, U_n'$ with order statistics

$$U_{1:2n} < \cdots < U_{2n:2n}.$$

Starting from the origin 0, move clockwise to the first of position of the pair $(U_1, U_1')$, continue clockwise to the first position of the pair $(U_2, U_2')$, and so on, continuing clockwise around the circle until the first of the two positions $(U_n, U_n')$ is encountered at a random index $1 \le I_n \le 2n$ (i.e. we stop at $U_{I_n:2n}$) after having made a random number $0 \le D_n \le n-1$ turns around the circle. Then for each positive integer $n$, the event $(I_n = 1)$ has probability

$$\mathbb{P}(I_n = 1) = \frac{1 - 2^n b_n(0)}{2n}$$

where $n! b_n(0) = B_n(0)$ is the $n$th Bernoulli number. For $1 \le k \le 2n$, the difference

$$\delta_{k:2n} := \frac{1}{2n} - \mathbb{P}(I_n = k)$$

is a polynomial function of $k$, which is closely related to $b_n(x)$. In particular, this difference has the surprising symmetry

$$\delta_{2n+1-k:2n} = (-1)^n \delta_{k:2n}, \quad \text{for } 1 \le k \le 2n$$

which is a combinatorial analog of the reflection symmetry (4.11) for the Bernoulli polynomials.

Stripping down the clock model, the random variables $I_n$ and $D_n$ are two statistics of permutations of the multiset

$$1^2 \cdots n^2 := \{1,1,2,2,\ldots,n,n\}. \tag{4.13}$$

Section 4.4 discusses the combinatorics behind the distributions of $I_n$ and $D_n$. In Section 4.5 we generalize the Bernoulli clock model to offer a new perspective on the work of Horton and Kurn [91] and the more recent work of Clifton et al [33]. In particular, we provide a probabilistic interpretation for the permutation counting problem in [91] and explicitly compute the mean function on $[0,1]$ of a renewal process with i.i.d. beta$(1,m)$-jumps. The expression of this mean function is given in terms of the complex roots of the exponential polynomial $E_m(x) := 1 + x/1! + \cdots + x^m/m!$, and its derivatives at 0 are precisely the moments of these roots, as studied in [168].

The circular convolution identities for Bernoulli polynomials are closely related to the decomposition of a real valued random variable $X$ into its integer part $\lfloor X \rfloor \in \mathbb{Z}$ and its fractional part $X^\circ \in \mathbb{T} := \mathbb{R}/\mathbb{Z} = [0,1)$:

$$X = \lfloor X \rfloor + X^\circ. \tag{4.14}$$

If $\gamma_1$ is a random variable with standard exponential distribution, then for each positive real $\lambda$ we have the expansion

$$\frac{d}{du}\mathbb{P}((\gamma_1/\lambda)^\circ \le u) = \frac{\lambda e^{-\lambda u}}{1-e^{-\lambda}} = B(u,-\lambda) = \sum_{n\ge 0} b_n(u)(-\lambda)^n. \tag{4.15}$$

Here the first two equalities hold for all real $\lambda \neq 0$ and $u \in [0,1)$, but the final equality holds with a convergent power series only for $0 < |\lambda| < 2\pi$. Section 4.6 presents a generalization of formula (4.15) with the standard exponential variable $\gamma_1$ replaced by the gamma distributed sum $\gamma_r$ of $r$ independent copies of $\gamma_1$, for a positive integer $r$. This provides an elementary probabilistic interpretation and proof of a formula due Erdélyi et al. [59, Section 1.11, page 30] relating the *Hurwitz-Lerch zeta function*

$$\Phi(z,s,u) = \sum_{m\ge 0} \frac{z^m}{(u+m)^s} \tag{4.16}$$

to Bernoulli polynomials.

## 4.2 Circular convolution of polynomials

Theorem 4.1 follows easily by induction on $n$ from the characterization (4.4) of the Bernoulli polynomials, and the action of circular convolution by the function

$$-b_1(u) = 1/2 - u, \tag{4.17}$$

as described by the following lemma.

**Lemma 4.3.** *For each Riemann integrable function $f$ with domain $[0,1)$, the circular convolution $h = f \circledast (-b_1)$ is continuous on $\mathbb{T}$, implying $h(0) = h(1-)$. Moreover,*

$$\int_0^1 h(u)du = 0 \tag{4.18}$$

*and at each $u \in (0,1)$ at which $f$ is continuous, $h$ is differentiable with*

$$\frac{d}{du}h(u) = f(u) - \int_0^1 f(v)dv. \tag{4.19}$$

*In particular, if $f$ is bounded and continuous on $(0,1)$, then $h = f \circledast (-b_1)$ is the unique continuous function $h$ on $\mathbb{T}$ subject to (4.18) with derivative (4.19) at every $u \in (0,1)$.*

*Proof.* According to the definition of circular convolution (4.8),

$$(f \circledast g)(u) = \int_0^u f(v)g(u-v)dv + \int_u^1 f(v)g(1+u-v)dv.$$

In particular, for $g(u) = -b_1(u)$, and a generic integrable function $f$,

$$\begin{aligned}(f \circledast (-b_1))(u) &= \int_0^u f(v)(v-u+1/2)dv + \int_u^1 f(v)(v-u-1/2)dv \\ &= \frac{1}{2}\left[\int_0^u f(v)dv - \int_u^1 f(v)dv\right] - u\int_0^1 f(v)dv + \int_0^1 vf(v)dv.\end{aligned}$$

Differentiate this identity with respect to $u$ to see that $h := f \circledast (-b_1)$ has the derivative displayed in (4.19) at every $u \in (0,1)$ at which $f$ is continuous, by the fundamental theorem of calculus. Also, this identity shows $h$ is continuous on $(0,1)$ with $h(0) = h(0+) = h(1-)$, hence $h$ is continous with respect to the topology of the circle $\mathbb{T}$. This $h$ has integral 0 by associativity of circular convolution: $h \circledast 1 = f \circledast (-b_1) \circledast 1 = f \circledast 0 = 0$. Assuming further that $f$ is bounded and continuous on $(0,1)$, the uniqueness of $h$ is obvious. □

The reformulation of Theorem 4.1 in Corollary 4.2 displays how simple it is to convolve Bernoulli polynomials on the circle. On the other hand, convolving monomials is less pleasant, as the following calculations show.

**Lemma 4.4.** *For real parameters $n > 0$ and $m > -1$,*

$$x^m \circledast x^n = x^n \circledast x^m = \frac{n}{m+1}x^{n-1} \circledast x^{m+1} + \frac{x^n - x^{m+1}}{m+1}. \tag{4.20}$$

*Proof.* Integrate by parts to obtain

$$\begin{aligned}x^n \circledast x^m &= \int_0^x u^n(x-u)^m du + \int_x^1 u^n(1+x-u)^m du \\ &= \frac{n}{m+1}\int_0^x u^{n-1}(x-u)^{m+1}du + \frac{n}{m+1}\int_x^1 u^{n-1}(1+x-u)^m du + \frac{x^n - x^{m+1}}{m+1}\end{aligned}$$

and hence (4.20). □

**Proposition 4.5** (Convolving monomials)**.** *For each positive integer $n$*

$$1 \circledast x^n = x^n \circledast 1 = \frac{1}{n+1}, \tag{4.21}$$

*and for all positive integers $m$ and $n$*

$$x^m \circledast x^n = x^n \circledast x^m = \frac{n!\ m!}{(n+m+1)!} + \sum_{k=0}^{n-1} \frac{n!}{(n-k)!(m+1)_{k+1}}(x^{n-k} - x^{m+k+1}) \tag{4.22}$$

*and with the Pochhammer notation $(m+1)_{k+1} := (m+1)\dots(m+k+1)$. In particular*

$$x \circledast x^n = \frac{x - x^{n+1}}{n+1} + \frac{1}{(n+1)(n+2)}.$$

*Proof.* By induction, using Lemma 4.4. □

**Remark 4.6.** 1. By inspection of (4.22) the polynomial $\left(x^n \circledast x^m - \frac{n!\ m!}{(n+m+1)!}\right)/x$ is an anti-reciprocal polynomial with rational coefficients.

2. Theorem 4.1 can be proved by inductive application of Proposition 4.5 to the expansion of the Bernoulli polynomials $B_n(x)$ in the monomial basis. This argument is unnecessarily complicated, but boils down to the two following identities for the Bernoulli numbers $B_n := B_n(0)$ for $n \geq 1$:

$$B_n = \frac{-1}{n+1} \sum_{k=0}^{n-1} \binom{n+1}{k} B_k \tag{4.23}$$

$$\frac{B_{n+1}}{(n+1)!} = \sum_{k=0}^{n} \frac{1}{(k+2)k!} \frac{B_{n-k}}{(n-k)!} \tag{4.24}$$

The identity (4.23) is a commonly used recursion for the Bernoulli numbers. We do not know any reference for (4.24), but this can be checked by manipulation of Euler's generating function (4.3).

3. Using the hypergeometric function $F := {}_2F_1$, it follows from Equation (4.22) that:

$$x^n \circledast x^m = \frac{n!m!}{(m+n+1)!} x^{m+n+1} + \frac{x^n}{m+1} F\left(1, -n; m+2; \frac{-1}{x}\right) - \frac{x^{m+1}}{m+1} F(1, -n; m+2; -x).$$

## 4.3 Probabilistic interpretation

For positive real numbers $a, b > 0$, recall that the beta$(a, b)$ probability distribution, has density

$$\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\ x^{a-1}(1-x)^{b-1}, \quad (0 \leq x \leq 1)$$

with respect the the Lebesgue measure on $\mathbb{R}$, where $\Gamma$ denotes Euler's gamma function. The following Corollary 4.7 offers a probabilistic interpretations of Theorem 4.1 in terms of summing i.i.d beta$(1, 2)$-distributed random variables on the circle.

**Corollary 4.7.** *The probability density of the sum of $n$ independent beta$(1, 2)$ random variables in the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is*

$$(1 - 2b_1)^{\circledast n}(u) = 1 - 2^n b_n(u), \quad \text{for } u \in \mathbb{T} = [0, 1).$$

*Proof.* Follows from Corollary 4.2. □

Recall that a beta$(1,2)$ random variable can be constructed as the minimum of two independent uniform random variables in $[0,1]$. Let $U_1, U_1', \ldots, U_n, U_n'$ be a sequence of $2n$ i.i.d random random variables with uniform distribution on $\mathbb{T} = [0,1)$. We think of these variables as random positions around a circle of circumference 1. On the event of probability one that the $U_i$ and $U_i'$ are all distinct, we define the following variables:

1. $U_{1:2n} < U_{2:2n} < \cdots < U_{2n:2n}$ the order statistics of $U_1, U_1', \ldots, U_n, U_n'$,
2. $X_1 := \min(U_1, U_1')$
3. for $2 \leq k \leq n$, the variable $X_k$ is the spacing around the circle from $X_{k-1}$ to whichever of $U_k, U_k'$ is encountered first moving cyclically around $\mathbb{T}$ from $X_{k-1}$,
4. $I_k$ is the random index in $\{1, \ldots, 2n\}$ such that $X_k = U_{I_k:2n}$.
5. $D_n \in \{0, \ldots, n-1\}$ is the random number of full rotations around $\mathbb{T}$ to find $X_n$. This is also the number of descents in the sequence $(I_1, I_2, \ldots, I_n)$; that is

$$D_n = \sum_{i=1}^{n-1} 1[I_i > I_{i+1}]. \tag{4.25}$$

We refer to this construction as the *Bernoulli clock*. Figure 4.1 depicts an instance of the Bernoulli clock for $n = 4$.

**Example 4.8.** In Figure 4.1, the clock is a circle of circumference 1. Inside the circle, the numbers $1, 2, \ldots, 8$ index the order statistics of 8 uniformly distributed random points on the circle. The corresponding numbers outside the circle are a random assignment of labels from the multiset of four pairs $1^2 2^2 3^2 4^2$. The four successive arrows delimit segments of $\mathbb{T} \equiv [0,1)$ whose lengths $X_1, X_2, X_3, X_4$ are independent beta$(1,2)$ random variables, while $(I_1, I_2, I_3, I_4)$ is the sequence of indices inside circle, at the end points of these four arrows. In this example, $(I_1, I_2, I_3, I_4) = (1, 4, 6, 3)$, and the number of turns around the circle is $D_4 = 1$.

**Proposition 4.9.** *With the above notation, the following hold*

1. *The random spacings $X_1, X_2, \ldots, X_n$ (defined by the Bernoulli clock above) are i.i.d beta$(1,2)$ random variables.*
2. *The random sequence of indices $(I_1, I_2, \ldots, I_n)$ is independent of the sequence of order statistics $(U_{1:2n}, \ldots, U_{2n:2n})$.*

*Proof.* Notice that $X_1 = \min(U_1, U_1')$ is a beta(1,2) random variable. Also, since $U_2, U_2'$ are i.i.d uniform and are independent of the random position of $X_1$ on the circle, the variable $X_2$ is independent from $X_1$ and also has distribution beta(1,2). With a similar argument we deduce that the variables $X_1, X_2 \ldots, X_n$ are i.i.d beta(1,2). Also, the random index $I_n$ at which the process stops depends only on the relative positions of the $U_i$'s and $U_i'$ 's. We then deduce that $I_n$ is independent of the order statistics $(U_{1:2n} < U_{2:2n} < \cdots < U_{2n:2n})$. □

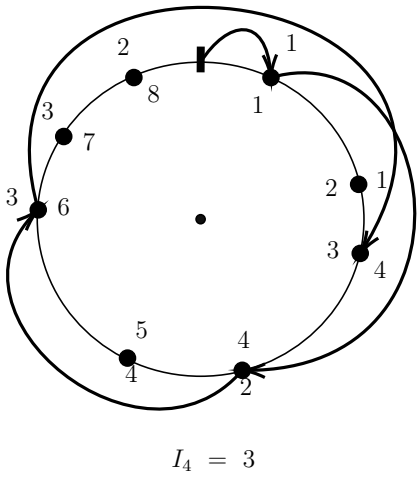



Figure 4.1: An instance of the Bernoulli clock model.

### 4.3.1 Expanding Bernoulli polynomials in the Bernstein basis

It is well known that, for $1 \le k \le 2n$, the distribution of $U_{k:2n}$ is beta$(k, 2n+1-k)$, whose probability density relative to Lebesgue measure at $u \in [0,1)$ is the normalized Bernstein polynomial of degree $2n-1$:

$$f_{k:2n}(u) := \frac{(2n)!}{(k-1)!(2n-k)!} u^{k-1}(1-u)^{2n-k}$$

**Proposition 4.10.** *For each positive integer $n$, the sum $S_n$ of $n$ independent beta$(1,2)$ variables has fractional part $S_n^\circ$ whose probability density on $(0,1)$ is given by the formulas*

$$f^\circ_{S_n}(u) = 1 - 2^n b_n(u) = \sum_{k=1}^{2n} p_{k:2n}\ f_{k:2n}(u), \quad \text{for } u \in (0,1). \tag{4.26}$$

*Here $(p_{1:2n}, \ldots, p_{2n:2n})$ is the probability distribution of the random index $I_n$ in the Bernoulli clock construction:*

$$p_{k:2n} = \mathbb{P}(I_n = k), \quad \text{for } 1 \le k \le 2n.$$

*Proof.* The first formula for the density of $S_n^\circ$ is read from Corollary 4.7. Proposition 4.9 represents $S_n^\circ = U_{I_n:2n}$ where the index $I_n$ is independent of the sequence of order statistics $(U_{k:2n}, 1 \le k \le 2n)$, hence the second formula for the same probability density on $(0,1)$. □

**Corollary 4.11.** *The factorially normalized Bernoulli polynomial of degree $n$ admits the expansion in Bernstein polynomials of degree $2n-1$*

$$b_n(u) = \frac{1}{2^n} \sum_{k=1}^{2n} \delta_{k:2n} \; f_{k:2n}(u) \tag{4.27}$$

*where $\delta_{k:2n}$ is the difference at $k$ between the uniform probability distribution on $\{1, \ldots, 2n\}$ and the distribution of $I_n$.*

$$\delta_{k:2n} = \frac{1}{2n} - p_{k:2n} \quad \textit{for } 1 \le k \le 2n. \tag{4.28}$$

*Proof.* Formula (4.27) is obtained from (4.26), in the first instance as an identity of continuous functions of $u \in (0,1)$, then as an identity of polynomials in $u$, by virtue of the binomial expansion

$$\sum_{k=1}^{2n} \frac{1}{2n} f_{k:2n}(u) = 1. \qquad \square$$

**Remark 4.12.** Since $b_n(1-u) = (-1)^n b_n(u)$ and $f_{k:2n}(1-u) = f_{2n+1-k:2n}(u)$, the identity (4.27) implies that the difference between the distribution of $I_n$ and the uniform distribution on $\{1, \ldots, 2n\}$ has the symmetry

$$\delta_{2n+1-k:2n} = (-1)^n \delta_{k:2n} \quad \text{for } 1 \le k \le 2n. \tag{4.29}$$

**Conjecture 4.13.** We conjecture that the discrete sequence $(\delta_{1:2n}, \ldots, \delta_{2n:2n})$ approximates the Bernoulli polynomials $b_n$ (hence also the shifted cosine functions, see (4.7)) as $n$ becomes large, more precisely:

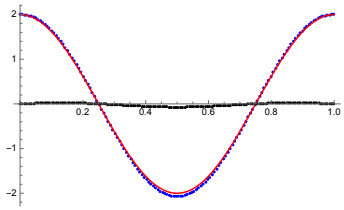
$$\sup_{1 \le k \le 2n} \left| 2n\pi^n \delta_{k:2n} - (2\pi)^n b_n\left(\frac{k-1}{2n-1}\right) \right| \to 0 \quad \text{as } n \to \infty.$$

Figure 4.3 does suggest that the difference $2n\pi^n \delta_n(k) - (2\pi)^n b_n\left(\frac{k-1}{2n-1}\right)$ gets smaller uniformly in $1 \le k \le 2n$ as $n$ grows, geometrically but rather slowly, like $C\rho^n$ for a constant $C > 0$ and $\rho \approx 2^{-1/100}$.

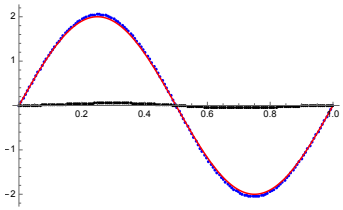
From (4.26) we see that we can expand the polynomial density $1 - 2^n b_n(u)$ in the Bernstein basis of degree $2n-1$ with positive coefficients. A similar expansion can obviously be achieved using Bernstein polynomials of degree $n$, with coefficients which must add to 1. These coefficients are easily calculated for modest values of $n$ (see (4.32)) which suggests the following

**Conjecture 4.14.** For each positive integer $n$, the polynomial probability density $1 - 2^n b_n(u)$ on $[0,1)$ can be expanded in the Bernstein basis of degree $n$ with positive coefficients.

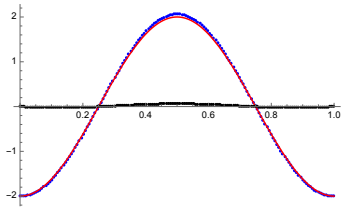
**Question 4.15.** More generally, what can be said about the greatest multiplier $c_n$ such that the polynomial $1 - c_n b_n(x)$ is a linear combination of degree $n$ Bernstein polynomials with non-negative coefficients?



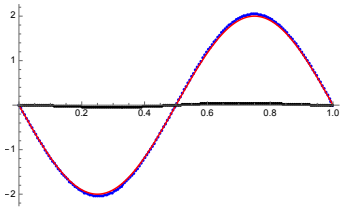
(a) $n = 70$



(b) $n = 75$



(c) $n = 80$



(d) $n = 85$

Figure 4.2: Plots of $2n\pi^n\delta_n$ (dotted curve in blue), $(2\pi)^n b_n(x)$ (curve in red) and their difference (dotted curve in black) for $n = 70, 75, 80, 85$.

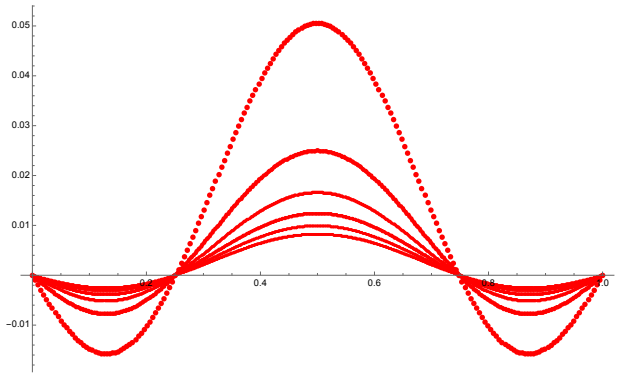


Figure 4.3: Plots of $2n\pi^n\delta_{k:2n} - (2\pi)^n b_n\left(\frac{k-1}{2n-1}\right)$ for $n = 100, 200, 300, 400, 500, 600$.

### 4.3.2 The distributions of $I_n$ and $D_n$

**Proposition 4.16.** *The distribution of $I_n$ in the Bernoulli clock construction is given by*

$$\mathbb{P}(I_n = k) = \frac{1}{2n} - \delta_{k:2n} \quad \textit{for } 1 \le k \le 2n \textit{ with} \tag{4.30}$$

$$\delta_{k:2n} = \frac{2^{n-1}}{n\ n!} \sum_{i=0}^{n} \frac{\binom{k-1}{i}\binom{n}{i}}{\binom{2n-1}{i}} B_{n-i}, \qquad \textit{for } 1 \le k \le 2n. \tag{4.31}$$

*Proof.* For each positive integer $N$, in the Bernstein basis $(f_{j:N})_{1\leq j\leq N}$ of polynomials of degree at most $N-1$, it is well known that the monomial $x^i$ can be expressed as

$$x^i = \frac{1}{N\binom{N-1}{i}} \sum_{j=i+1}^{N} \binom{j-1}{i} f_{j:N}(x) \quad \text{for } 0 \leq i < N.$$

Plugging this expansion into (4.2) yields the expansion of $b_n(x)$ in the Bernstein basis of degree $N-1$ for every $N > n$

$$b_n(x) = \sum_{j=1}^{N} \left( \sum_{i=0}^{n} \frac{\binom{j-1}{i}\binom{n}{i}}{n!N\binom{N-1}{i}} B_{n-i} \right) f_{j:N}(x) \qquad (0 \leq n < N). \tag{4.32}$$

In particular, for $N = 2n$ comparison of this formula with (4.27) yields (4.31) and hence (4.30) □

**Remark 4.17.** The error $\delta_{k:2n}$ is polynomial in $k$ and the symmetry $\delta_{2n+1-j:2n} = (-1)^n \delta_{j:2n}$ is not obvious from (4.31).

Let us now derive the distribution of $D_n$ explicitly. From the Bernoulli clock scheme, we can construct the random variable $D_n$ as follows. Let $X_1, \ldots, X_n$ be a sequence of i.i.d random variables and $S_n := X_1 + \cdots + X_n$ their sum in $\mathbb{R}$ (not in the circle $\mathbb{T}$). We have

$$D_n = \lfloor S_n \rfloor.$$

**Theorem 4.18.** *The distribution function of $S_n$ is given by*

$$\mathbb{P}(S_n \leq x) = 2^n \sum_{k=0}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^{n-k-j} \frac{(x-k)_+^{2n-j}}{(2n-j)!}, \qquad \textit{for } x \geq 0,$$

*where $x_+$ denotes* $\max(x, 0)$ *for $x \in \mathbb{R}$.*

*Proof.* Let $\varphi$ be the Laplace transform of the $X_i$'s i.e.

$$\varphi_X(\theta) := \mathbb{E}[e^{-\theta X_1}] = \int_0^{+\infty} \theta e^{-\theta t} \mathbb{P}(X_1 \leq t) dt.$$

We compute $\varphi_X$ and we obtain

$$\varphi_X(\theta) = \frac{2}{\theta^2} \left( e^{-\theta} + (\theta - 1) \right).$$

So for $n \geq 1$, the Laplace transform of $S_n$ is then given by

$$\varphi_{S_n}(\theta) = (\varphi_X(\theta))^n = 2^n \sum_{k=0}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^{n-k-j} \frac{e^{-k\theta}}{\theta^{2n-j}}. \tag{4.33}$$

The transform $\varphi_{S_n}$ can be inverted term by term using the following identity

$$\int_0^{+\infty} \theta e^{-\theta t} \frac{(t-k)_+^n}{n!}\, dt = \frac{e^{-k\theta}}{\theta^n}, \quad \text{for } k \geq 0,\ \theta > 0 \text{ and } n \geq 0. \tag{4.34}$$

We then obtain the cdf of $S_n$ as follows:

$$\mathbb{P}(S_n \leq x) = 2^n \sum_{k=0}^{n} \sum_{j=0}^{n-k} \binom{n}{k} \binom{n-k}{j} (-1)^{n-k-j} \frac{(x-k)_+^{2n-j}}{(2n-j)!}, \quad \text{for } x \geq 0. \tag{4.35}$$

□

**Remark 4.19.** (4.34) was known to Lagrange in the 1700s and it appears in [106, Lemme III and Corollaire I] where he said the final words on inverting Laplace transforms of the form (4.33):

> *"... mais comme cette intégration est facile par les methodes connues, nous n'entrerons pas dans un plus grand detail là-dessus; et nous terminerons même ici nos recherches, par lesquelles on doit voir qu'il ne reste plus de difficulté dans la solution des questions qu'on peut proposer à ce sujet."*

Since $S_n$ has a density, we can deduce that

$$\mathbb{P}(D_n = k) = \mathbb{P}(S_n \leq k+1) - \mathbb{P}(S_n \leq k), \quad \text{for } 0 \leq k \leq n-1.$$

Combined with (4.35) this gives the distribution of $D_n$ explicitly. The following table gives $\#(n;+,\bullet)$ for small values of $n$ and $d$.

| $n$ \ $d$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 1 | 1 | | | | | |
| 2 | 1 | 1 | | | | |
| 3 | 47 | 42 | 1 | | | |
| 4 | 641 | 1659 | 219 | 1 | | |
| 5 | 11389 | 72572 | 28470 | 968 | 1 | |
| 6 | 248749 | 3610485 | 3263402 | 357746 | 4017 | 1 |

Table 4.1: The table of $\#(n;+,d)$.

**Remark 4.20.** The sequence $a(n) = \#(n;+,0) = 2^{-n}(2n)!\ \mathbb{P}(D_n = 0)$, which counts the number of permutations of $1^2 \cdots n^2$ for which $D_n = 0$ (the first column in Table 4.1), can be explicitly written using (4.35) as follows

$$a(n) = \mathbb{P}(S_n \leq 1) = \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} \frac{(2n)!}{(2n-j)!}. \tag{4.36}$$

This integer sequence appears in many other contexts (see OEIS entry A006902), among which we mention a few:

1. $a(n)$ is the number of words on $1^2 \cdots n^2$ with longest complete increasing sub-sequence of length $n$. We shall detail this in Section 4.5.

2. $a(n) = n!\, Z(\mathfrak{S}_n; n, n-1, \ldots, 1)$ where $Z(\mathfrak{S}_n)$ is the cycle index of the symmetric group of order $n$ (see [153, Section 1.3]).

3. $a(n) = \mathrm{B}_n\left(n \cdot 0!,\ (n-1) \cdot 1!,\ (n-2)! \cdot 2!,\ \ldots,\ 1 \cdot (n-1)!\right)$, where $\mathrm{B}_n(x_1, \ldots, x_n)$ is the $n$-th complete Bell polynomial.

## 4.4 Combinatorics of the Bernoulli clock

There are a number of known constructions of the Bernoulli numbers $B_n$ by permutation enumerations. Entringer [58] showed that Euler's presentations of the Bernoulli numbers, as coefficients in the expansions of hyperbolic and trigonometric functions, lead to explicit formulas for $B_n$ by enumeration of alternating permutations. More recently, Graham and Zang [82] gave a formula for $B_{2n}$ by enumerating a particular subset of the set of $2^{-n}(2n)!$ permutations of the multiset $1^2 \cdots n^2$ of $n$ pairs.

The number of permutations of this multiset, such that for every $i < n$ between each pair of occurrences of $i$ there is exactly one $i+1$, is $(-2)^n(1-2^{2n})B_{2n}$. Here we offer a novel combinatorial expression of the Bernoulli numbers based on a different attribute of permutations of same multiset (4.13), which arises from the the probabilistic interpretation in Section 4.3. We call the combinatorial construction involved the *the Bernoulli clock*. Fix a positive integer $n \geq 1$ and for a permutation $\tau$ of the multiset (4.13),

- Let $1 \leq I_1 \leq 2n-1$ be the position of the first 1; that is $I_1 = \min\{1 \leq k \leq 2n \colon \tau(k) = 1\}$.
- For $1 \leq k \leq n-1$, denote by $1 \leq I_{k+1} \leq 2n$ the index of the first value $k+1$ following $I_k$ in the cyclic order (circling back to the beginning of necessary).
- Let $0 \leq D_n \leq n-1$ be the number of times we circled back to the beginning of the multiset before obtaining the last index $I_n$.

**Example 4.21.** The permutation $\tau$ corresponding to Figure 4.1 is the permutation $\tau = (1, 1, 4, 2, 4, 3, 3, 2)$. For this permutation

$$(I_1, I_2, I_3, I_4) = (1, 4, 6, 3) \quad \text{and} \quad D_4 = 1.$$

Notice that random index $I_n$ and the number of descents $D_n$ depend only on the relative positions of $U_1, U_1', \ldots, U_n, U_n'$ i.e. the permutation of the multiset $1^2 \cdots n^2$. So the distribution of $I_n$ and $D_n$ can be obtained by enumerating permutations. For $n \geq 1$, $1 \leq i \leq 2n$ and $0 \leq d \leq n-1$, let us denote by

1. $\#(n; i, d)$ the number of permutations among the $(2n)!/2^n$ permutations of the multiset $\{1, 1, \ldots, n, n\}$ that yield $I_n = i$ and $D_n = d$,
2. $\#(n; i, +)$ the number of permutations that yield $I_n = i$,
3. $\#(n; +, d)$ the number of permutations that yield $D_n = d$.

For $n = 2$ there are 6 permutations of $\{1, 1, 2, 2\}$ summarized in the following table

| Permutations | 1122 | 1212 | 1221 | 2112 | 2121 | 2211 |
|---|---|---|---|---|---|---|
| $(I_2, D_2)$ | $(3, 0)$ | $(2, 0)$ | $(2, 0)$ | $(4, 0)$ | $(3, 0)$ | $(1, 1)$ |

Table 4.2: Permutations of $\{1, 1, 2, 2\}$ and corresponding values of $(I_2, D_2)$.

The joint distribution of $I_2, D_2$ is then given by

| $I_2$ \ $D_2$ | 1 | 2 | 3 | 4 | $\#(2; +, \bullet)$ |
|---|---|---|---|---|---|
| 0 | 0 | 2 | 2 | 1 | 5 |
| 1 | 1 | 0 | 0 | 0 | 1 |
| $\#(2; \bullet, +)$ | 1 | 2 | 2 | 1 | 6 |

Table 4.3: The table of $\#(2; \bullet, \bullet)$.

Similarly for $n = 3$ we get

| $I_3$ \ $D_3$ | 1 | 2 | 3 | 4 | 5 | 6 | $\#(3; +, \bullet)$ |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 6 | 12 | 15 | 14 | 47 |
| 1 | 14 | 13 | 8 | 4 | 2 | 1 | 42 |
| 2 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| $\#(3; \bullet, +)$ | 15 | 13 | 14 | 16 | 17 | 15 | 90 |

Table 4.4: The table of $\#(3; \bullet, \bullet)$.

The distribution of $(I_n, D_n)$ can be obtained recursively as follows. The key observation is that every permutation of the multi-set $1^2 2^2 \cdots n^2$ is obtained by first choosing a permutation

of $1^2 2^2 \cdots (n-1)^2$, then choosing 2 places to insert the two values $n, n$. There are $\binom{2(n-1)}{2}$ options for where to insert the two last values. This corresponds to the factorization

$$(2n)!\ 2^{-n} = (2(n-1))!\ 2^{-n+1} \binom{2n}{2}.$$

Moreover, for $x \in \{1, \dots, 2(n-1)\}$ the identity of quadratic polynomials

$$\binom{x+1}{2} + \binom{2n-x}{2} + x(2n-1-x) = \binom{2n}{2},$$

translates, for each integer $x \in \{1, \dots, 2(n-1)\}$ and each permutation $\sigma$ of $1^2, \cdots (n-1)^2$, the decomposition of the total number of ways to insert the next two values $n, n$ according to whether:

1. both places are to the left of $x$,
2. both places are to the right of $x$,
3. one of those places is to the left of $x$ and the other to the right of $x$.

Suppose we ran the Bernoulli clock scheme on $2(n-1)$ hours and obtained $(I_{n-1}, D_{n-1})$. Inserting two new values $n, n$, the index $I_n$ then depends only on $I_{n-1}$ and the places where the two new values $n$ are inserted relatively to $I_{n-1}$. So, the sequence $(I_1, I_2, \dots)$ is a time-inhomogeneous Markov chain starting from $I_1 = 1$ and a $2(n-1) \times 2n$ transition matrix from $I_{n-1}$ to $I_n$ given by

$$P_n(x \to y) = \mathbb{P}(I_n = y | I_{n-1} = x) = \frac{Q_n(x,y)}{\binom{2n}{2}}, \qquad (1 \le x \le (2n-1),\ 1 \le y \le 2n)$$

where $Q_n(x,y)$ is the number of ways to insert the two new values $n$ in the Bernoulli clock in such a way that the first one of them to the right of $x$ is at place $y$. More explicitly, by elementary counting, we have

$$Q_n(x,y) = \begin{cases} x - y + 1, & \text{if } 1 \le y \le x \\ 2n - 1 - x, & \text{if } y = x+1 \\ 2n - y + x, & \text{if } x+2 \le y \le 2n \end{cases}$$

So the first few transition matrices are

$$P_2 = \frac{Q_2}{\binom{4}{2}} = \frac{1}{6} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix}, \qquad P_3 = \frac{Q_3}{\binom{6}{2}} = \frac{1}{15} \begin{pmatrix} 1 & 4 & 4 & 3 & 2 & 1 \\ 2 & 1 & 3 & 4 & 3 & 2 \\ 3 & 2 & 1 & 2 & 4 & 3 \\ 4 & 3 & 2 & 1 & 1 & 4 \end{pmatrix},$$

$$\text{and} \quad P_4 = \frac{Q_4}{\binom{8}{2}} = \frac{1}{28}\begin{pmatrix} 1 & 6 & 6 & 5 & 4 & 3 & 2 & 1 \\ 2 & 1 & 5 & 6 & 5 & 4 & 3 & 2 \\ 3 & 2 & 1 & 4 & 6 & 5 & 4 & 3 \\ 4 & 3 & 2 & 1 & 3 & 6 & 5 & 4 \\ 5 & 4 & 3 & 2 & 1 & 2 & 6 & 5 \\ 6 & 5 & 4 & 3 & 2 & 1 & 1 & 6 \end{pmatrix},$$

see Table 4.5 for a detailed combinatorial construction of $Q_3$. This discussion is summarized by the following proposition.

**Proposition 4.22.** *For a uniform random permutation of $1^2 \cdots n^2$ the probability distribution of $I_n$, treated as a $1 \times 2n$ row vector $p_n = (p_{1:2n}, \ldots, p_{2n:2n})$, is determined recursively by the matrix forward equations*

$$p_{n+1} = p_n \; P_{n+1} \qquad \textit{for } n = 1, 2, \ldots \textit{ starting from } p_1 = (1, 0). \tag{4.37}$$

So the first few of these distributions of $I_n$ are as follows:

$$p_1 = (1, 0), \qquad p_2 = \frac{1}{6}(1, 2, 2, 1),$$
$$p_3 = \frac{1}{90}(15, 13, 14, 16, 17, 15), \qquad p_4 = \frac{1}{2520}(322, 322, 312, 304, 304, 312, 322, 322).$$

As $n$ become bigger, the distribution $p_n$ gets closer to the uniform on $\{1, \ldots, 2n\}$. The error $\delta_n(k) = 1/(2n) - p_{k:2n}$ is polynomial in $k$ and satisfies the same forward equation as $p_n$ i.e.

$$\delta_{n+1} = \delta_n \; P_{n+1} \qquad \text{for } n = 1, 2, \ldots \text{ starting from } \delta_0 = (1/2, -1/2). \tag{4.38}$$

The sequence $\delta_n$ is also closely tied to the polynomial $b_n(x)$ as (4.27) shows.

**Example 4.23.** : The top $1 \times 6$ row of Table 4.5 displays the column index of places in rows of the main $15 \times 6$ table below it. The 15 rows of the main table list all $\binom{6}{2} = 15$ pairs of places, represented as two dots $\bullet$, in which two new values $3, 3$ can be inserted relative to 4 possible places of $I_2 \in \{1, 2, 3, 4\}$. The exponents of each dot $\bullet$ are the values of $I_2$ leading to $I_3$ being the column index of that dot in $\{1, 2, 3, 4, 5, 6\}$. For example in the second row, representing insertions of the new value 3 in places 1 and 3 of 6 places, the dot $\bullet^{2,3,4}$ in place 1 is the place $I_3$ found by the Bernoulli clock algorithm if $I_2 \in \{2, 3, 4\}$. The matrix $Q_3$ is the $4 \times 6$ matrix below the main table. The entry $Q_3(i, j)$ in row $i$ and column $j$ of $Q_3$ is the number of times $i$ appears in the exponent of a dot $\bullet$ in the $j$-th column of the main table.

**Remark 4.24.** Notice that the matrices $Q_n$ have the remarkable symmetry

$$2n - 1 - Q_n(i, j) = \widetilde{Q}_n(i, j), \quad (1 \le i \le 2n, \; 1 \le j \le 2n + 2), \tag{4.39}$$

with $\widetilde{Q}_n(i, j) \coloneqq Q_n(2n - 1 - i, 2n + 1 - j)$ i.e. the matrix $\widetilde{Q}_n$ is the matrix $Q_n$ with entries in reverse order in both axis.

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| $\bullet^{1,2,3,4}$ | $\bullet$ | 1 | 2 | 3 | 4 |
| $\bullet^{2,3,4}$ | 1 | $\bullet^{1}$ | 2 | 3 | 4 |
| $\bullet^{3,4}$ | 1 | 2 | $\bullet^{1,2}$ | 3 | 4 |
| $\bullet^{4}$ | 1 | 2 | 3 | $\bullet^{1,2,3}$ | 4 |
| $\bullet$ | 1 | 2 | 3 | 4 | $\bullet^{1,2,3,4}$ |
| 1 | $\bullet^{1,2,3,4}$ | $\bullet$ | 2 | 3 | 4 |
| 1 | $\bullet^{1,3,4}$ | 2 | $\bullet^{2}$ | 3 | 4 |
| 1 | $\bullet^{1,4}$ | 2 | 3 | $\bullet^{2,3}$ | 4 |
| 1 | $\bullet^{1}$ | 2 | 3 | 4 | $\bullet^{2,3,4}$ |
| 1 | 2 | $\bullet^{1,2,3,4}$ | $\bullet$ | 3 | 4 |
| 1 | 2 | $\bullet^{1,2,4}$ | 3 | $\bullet^{3}$ | 4 |
| 1 | 2 | $\bullet^{1,2}$ | 3 | 4 | $\bullet^{3,4}$ |
| 1 | 2 | 3 | $\bullet^{1,2,3,4}$ | $\bullet$ | 4 |
| 1 | 2 | 3 | $\bullet^{1,2,3}$ | 4 | $\bullet^{4}$ |
| 1 | 2 | 3 | 4 | $\bullet^{1,2,3,4}$ | $\bullet$ |

| | | | | | |
|---|---|---|---|---|---|
| **1** | **4** | **4** | **3** | **2** | **1** |
| **2** | **1** | **3** | **4** | **3** | **2** |
| **3** | **2** | **1** | **2** | **4** | **3** |
| **4** | **3** | **2** | **1** | **1** | **4** |

Table 4.5: The combinatorial construction of the matrix $Q_3$.

**Remark 4.25.**

1. It is interesting to note that, from (4.37), it is not clear what the Bernoulli polynomials have in relation with the distribution $p_n$ or the error $\delta_n$. It is not also clear from this recursion, even with (4.39), that $\delta_n$ has the symmetry described in (4.29).

2. Considering $\delta_n$ as a discrete analogue of $b_n$, one can think of the equation $\delta_{n+1} = \delta_n \, P_{n+1}$ as a discrete analogue of the integral formula (4.4).

3. In addition to the dynamics of the Markov chain $I = (I_1, I_2, \dots)$, we can get obtain the joint distribution of $(I_n, D_n)$ recursively in the same way. The key observation is that at step $n$, having obtained $I_n$ from the Bernoulli clock scheme and inserting the two new values $n+1$ in the clock, we either increment $D_n$ by 1 to get $D_{n+1}$ if both values are inserted prior to $I_n$ or the number of laps is not incremented i.e. $D_{n+1} = D_n$ if one of the two values is inserted after $I_n$. We then obtain the following recursion for $\#(n; i, d)$:

1) $\#(1;1,0) = 1$

2) $\#(n+1;i,d) = \sum_{1\le x<h} \#(n;i,x)\ \#_{n+1}(x,d) + \sum_{h\le x\le 2n} \#(n;i-1,x)\ \#_{n+1}(x,d).$

So one can get the joint distribution of $(I_n, D_n)$ recursively with

$$\mathbb{P}(I_n = i,\ D_n = d) = \frac{\#(n;i,d)}{2^{-n}(2n)!}.$$

## 4.5 Generalized Bernoulli clock

Let $n \ge 1$, $m_1, \dots, m_n \ge 1$ be positive integers and $M = m_1 + \cdots + m_n$. Let $\tau_n = \tau(m_1, \dots, m_n)$ be a random permutation uniformly distributed among the $M!/(m_1! \dots m_n!)$ permutations of the multiset $1^{m_1} 2^{m_2} \dots n^{m_n}$. Let us denote by $1 \le I_1 \le M$ the index of the first 1 in the sequence $\tau_n$. Continuing from this index $I_1$, let $I_2$ be the index of the first 2 we encounter (circling back if necessary) and continuing in this manner we get random indices $(I_1, I_2, \dots, I_n)$. Let us denote by $D_n = D(m_1, \dots, m_n)$ the number of times we circled around the sequence $\tau_n$ in this process, that is the number of descents in the random sequence $(I_1, I_2, \dots, I_n)$, as in (4.25).

For the continuous model, mark the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z} \cong [0,1)$ with $M$ i.i.d uniform on $[0,1]$ random variables $U_1^{(1)}, \dots, U_1^{(m_1)}$ , ... , $U_n^{(1)}, \dots, U_n^{(m_n)}$ and let $U_{1:M} < \cdots < U_{M:M}$ be their order statistics. Starting from 0 we walk around the clock until we encounter the first of the variables $U_1^{(i)}$ at some random index $I_1$. We continue from the random index $I_1$ until we encounter the first of the variables $U_2^{(i)}$ (circling back if necessary) and continue like this until we encounter the first of the variables $U_n^{(i)}$. We then obtain the random sequence $(I_1, I_2, \dots, I_n)$ and $D_n$ is the number of times we circled around the clock. Finally, let us denote by $(X_1, \dots, X_n)$ the lengths (clock-wise) of the segments $[U_{I_1:M}, U_{I_2:M}]$, $\cdots$, $[U_{I_{n-1}:M}, U_{I_n:M}]$, $[U_{I_n:M}, U_{I_1:M}]$ on the clock. The model described in Section 4.4 is the particular instance of this model where $m_1 = \cdots = m_n = 2$.

**Remark 4.26.** When there is no risk of confusion, we suppress the parameters $m_1, \cdots, m_n$ to simplify the notation.

**Proposition 4.27.** *The following hold:*

1. *The random lengths $X_1, X_2, \dots, X_n$ are independent random variables and $X_i$ has distribution beta$(1, m_i)$ for each $1 \le i \le n$.*

2. *The random sequence of indices $(I_1, I_2, \dots, I_n)$ is independent of the order statistics $(U_{1:M} < \cdots < U_{M:M})$.*

*Proof.* Notice that $X_1 = \min(U_1^{(1)}, \dots, U_1^{(m_1)})$ is a beta$(1, m_1)$ random variable. Also, since $U_2^{(1)}, \dots, U_2^{(m_2)}$ are i.i.d uniform and are independent of the position of $X_1$ on the circle,

the variables $U_2^{(i)} - X_1 \mod \mathbb{Z} \in [0,1)$ are still i.i.d uniform so $X_2$ is also beta$(1, m_2)$ and independent of $X_1$. Running the same argument repeatedly we deduce that the variables $X_1, X_2 \dots, X_n$ are independent with $X_i \sim \text{beta}(1, m_i)$. Also, the random index $I_n$ at which the process stops depends only on the relative positions of the variables $U_1^{(1)}, \dots, U_1^{(m_1)}, \dots, U_n^{(1)}, \dots, U_n^{(m_n)}$ i.e. $I_n$ is fully determined by the random permutation of $\{1, \dots, M\}$ induced by the $M$ i.i.d uniforms. We then deduce that $I_n$ is independent of the order statistics $(U_{1:M} < U_{2:M} < \cdots < U_{M:M})$. □

The number $D_n$ of turns around the clock can also be expressed as follows

$$D_n = \lfloor S_n \rfloor, \quad \text{where } S_n \coloneqq X_1 + \cdots + X_n. \tag{4.40}$$

Let us denote by $L_n = L(m_1, m_2, \dots, m_n)$ the length of the longest continuous increasing subsequence of $\tau_n$ starting with 1; that is the largest integer $1 \le \ell \le n$ such that

$$1, 2, 3, \dots, \ell \quad \text{is a subsequence of } \tau_n.$$

**Example 4.28.** Suppose $n = 4$ and $(m_1, m_2, m_3, m_4) = (2, 3, 2, 4)$ and consider the permutation $\tau_n = (\mathbf{1}, 4, 4, 1, 4, \mathbf{2}, 4, \mathbf{3}, 3, 2, 2)$. The longest increasing continuous subsequence of $\tau_n$ starting from 1 (the boldfaced subsequence) has length $L_4 = 3$ in this case.

For an infinite sequence $m = (m_1, m_2, \cdots)$ of positive integers, notice that we can construct the sequences of variables $L_n = L(m_1, \dots, m_n)$ , $D_n = D(m_1, \dots, m_n)$ and $I_n = I(m_1, \dots, m_n)$ on a common probability space. This is done by marking an additional $m_n$ i.i.d uniform positions on the circle $\mathbb{T}$ at each step $n$. Notice then that $(L_n = L(m_1, \dots, m_n))_{n \ge 1}$ is an increasing sequence of random variables so we define

$$L_\infty \coloneqq \lim_{n \to \infty} L_n \quad \text{and} \quad \mathcal{L}_m \coloneqq \mathbb{E}[L_\infty].$$

**Proposition 4.29.** *We have the following:*

$$L_n = \sum_{k=0}^{n} 1[S_k \le 1] \quad \textit{and} \quad L_\infty = \sum_{k=0}^{\infty} 1[S_k \le 1].$$

*In particular, we have* $(L_n = n) = (D_n = 0)$ *and for* $n \ge k$ *we have*

$$(L(m_1, \dots, m_n) \ge k) = (L(m_1, \dots, m_k) = k).$$

*Proof.* The length $L_n$ of the longest sequence of the form $1 \dots \ell$ is the maximal integer $\ell$ such that $S_\ell \le 1$, i.e. the maximal $l$ such that the random walk $(S_k)_{k \ge 0}$ does not shoot over 1. Then we deduce that indeed

$$L_n = \sum_{k=0}^{n} 1[S_k \le 1].$$

The rest of the statements follow immediately from this equation. □

**Corollary 4.30.** *For $k \leq n$ we have*

$$\mathbb{P}(L_n \geq k) = \mathbb{P}(S_k \leq 1).$$

*Proof.* Follows immediately from Proposition 4.29. □

**Remark 4.31.** When $m_1 = m_2 = \dots m_n = 1$, the random variable $S_n$ is the sum of $n$ i.i.d uniform random variables on $[0,1]$ and the fractional part $S_n^\circ$ has uniform distribution on $\mathbb{T}$. The index $I_n$ has uniform distribution in $\{1,\dots,n\}$ and the distribution of the number of descents

$$P(D_n = k) = \frac{A_{n,k}}{n!}, \quad (0 \leq k \leq n-1)$$

is given by the Eulerian numbers $A_{n,k}$, see [153, Section 1.4].

Horton and Kurn [91, Theorem and Corollary (c)] gives a formula for the number of permutations $\tau$ of the multiset $1^{m_1}2^{m_2}\dots n^{m_n}$ for which $L_n = n$; that is a formula for

$$\frac{M!}{m_1!\cdots m_n!}\,\mathbb{P}(L_n = n).$$

We shall interpret this formula in our context and rederive it from a probabilistic perspective.

**Theorem 4.32.** *The number of permutations $\tau_n$ of the multiset $1^{m_1}2^{m_2}\dots n^{m_n}$ that contain the sequence $(1,2,\cdots,n)$ is given by*

$$\frac{M!}{m_1!\cdots m_n!}\mathbb{P}(L_n = n) = (-1)^M \sum_{j=0}^{M} \binom{M}{j} \frac{c_j}{j!}, \tag{4.41}$$

*where*

$$c_j = (-1)^n [\theta^j] \prod_{i=1}^{n} E_{m_i - 1}(-\theta),$$

*with $[x^n]f(x)$ denoting the coefficient of $x^n$ in the power series expansion of $f$.*

*Proof.* Similarly to our discussion in Section 4.4, we can obtain an expression for $\mathbb{P}(S_n \leq x)$ by inverting the Laplace transform of $S_n$. First recall that the Laplace transform of $X_i \sim \text{beta}(1, m_i)$ is

$$\varphi_{X_i}(\theta) = \mathbb{E}[e^{-\theta X_i}] = (-1)^{m_i} \frac{m_i!}{\theta^{m_i}} \left(e^{-\theta} - E_{m_i-1}(-\theta)\right),$$

where $E_k(x)$ denotes the exponential polynomial $E_k(x) = \sum_{i=0}^{k} x^i/i!$. So the Laplace transform of $S_n$ is then given by

$$\varphi_{S_n}(\theta) = \frac{(-1)^M \prod_{i=1}^{n} m_i!}{\theta^M} \prod_{i=1}^{n} \left(e^{-\theta} - E_{m_i-1}(-\theta)\right).$$

Let us write the following product as a polynomial in two variables $X$ and $\theta$:

$$\prod_{i=1}^{n}(X - E_{m_i-1}(-\theta)) = \sum_{k,j\geq 0}\alpha_{k,j}\theta^j X^k, \tag{4.42}$$

so that

$$\varphi_{S_n}(\theta) = (-1)^M\left(\prod_{i=1}^{n} m_i!\right)\sum_{k,j\geq 0}\alpha_{k,j}\frac{e^{-\theta}}{\theta^{M-j}}.$$

Using (4.34) to invert this Laplace transform, we get

$$\mathbb{P}(S_n \leq x) = (-1)^M\left(\prod_{i=1}^{n} m_i!\right)\sum_{k,j\geq 0}\alpha_{k,j}\frac{(x-k)_+^{M-j}}{(M-j)!},$$

So we deduce that

$$\mathbb{P}(L_n = n) = \mathbb{P}(S_n \leq 1) = (-1)^M\left(\prod_{i=1}^{n} m_i!\right)\sum_{j=0}^{M}\frac{c_j}{(M-j)!},$$

where, from (4.42) we have

$$c_j = \alpha_{0,j} = (-1)^n[\theta^j]\left(\prod_{i=1}^{n} E_{m_i-1}(-\theta)\right).$$

Multiplying by $M!/(m_1!\cdots m_n!)$ we get the formula in (4.41). □

We suppose from now on that $m := m_1 = m_2 = \cdots \geq 1$. Let $\mathcal{L}_{n,m}$ and $\mathcal{L}_m$ denote the expectation of $L_n$ and $L_\infty$; that is

$$\mathcal{L}_{n,m} := \mathbb{E}[L_n] \quad \text{and} \quad \mathcal{L}_m := \lim_{n\to\infty}\mathcal{L}_{n,m} = \mathbb{E}[L_\infty].$$

In [33], Clifton et al. give a fine asymptotic study of $\mathcal{L}_m$ as $m \to \infty$. In this paper, we provide a pleasant probabilistic framework in which the discussion [33] fits rather naturally.

Let $(N(t), t \geq 0)$ be the renewal process with beta$(1, m)$-distributed i.i.d jumps $X_i$ i.e.

$$N(t) = \sum_{n\geq 1}^{\infty} 1[S_n \leq t].$$

Notice that, by virtue of Proposition 4.29, the variable $N(1) = L_\infty - 1$ is the number of renewals of $N$ in $[0, 1]$. Let $M(t) := \mathbb{E}[N(t)]$ denote the mean of $N(t)$. By first step analysis, $M(t)$ satisfies the following equation for $t \in [0, 1]$:

$$M(t) = \mathbb{P}(X_1 \leq t) + m \int_0^t M(t-x)(1-x)^{m-1} dx, \tag{4.43}$$
$$= \mathbb{P}(X_1 \leq t) + m \int_0^t M(x)(1-t+x)^{m-1} dx.$$

From (4.43) we can deduce that $M$ satisfies the following differential equation

$$1 + \sum_{k=0}^{m} \frac{(-1)^k}{k!} M^{(k)}(t) = 0. \tag{4.44}$$

**Theorem 4.33.** *Let $\alpha_1, \ldots, \alpha_m$ be the $m$ distinct complex roots of the exponential polynomial $E_m(x) = \sum_{k=0}^{m} x^k/k!$. Then the mean function $M(t)$ is given by*

$$M(t) = -1 - \sum_{k=1}^{m} \alpha_k^{-1} e^{-\alpha_k t}. \tag{4.45}$$

Before we prove Theorem 4.33, we first recall a couple of intermediate results.

**Lemma 4.34.** *Let $z$ be a non-zero complex number. Then, for any positive integer $n$ and $t \in [0,1]$, we have the following:*

$$\int_0^t e^{zx}(1-x)^n dx = n! \sum_{j=0}^{n} \frac{e^{zt}(1-t)^j - 1}{j!} z^{j-n-1}.$$

*Proof.* Follows immediately by induction on $n$ and integration by parts. □

The following lemma is an adaptation of [168, Theorem 7].

**Lemma 4.35.** *Let $\alpha_1, \ldots, \alpha_m$ be the $m$ distinct complex zeros of $E_m(x)$. Then we have the following*

$$\sum_{k=1}^{m} \alpha_k^{-j} = \begin{cases} -1, & \text{if } j = 1, \\ 0, & \text{if } 2 \leq j \leq m, \\ 1/m!, & \text{if } j = m+1. \end{cases}$$

*Proof of Theorem 4.33.* The mean function $M(t)$ satisfies (4.44). The latter is an order $m$ ODE with constant coefficients and its characteristic polynomial is $E_m(-x)$ whose roots are $-\alpha_1, \ldots, -\alpha_m$. So the solution is of the form

$$M(t) = -1 + \sum_{k=1}^{m} \beta_k e^{-\alpha_k t}.$$

Setting $\beta_k = -\alpha_k^{-1}$ for $1 \leq k \leq m$, it suffices to show that $M(t)$ satisfies (4.43). To that end notice that, thanks to Lemma 4.34, we have

$$\begin{aligned}
&\mathbb{P}(X_1 \leq t) + m \int_0^t M(t-x)(1-x)^{m-1} dx \\
&= m \int_0^t (1 + M(t-x))(1-x)^{m-1} dx \\
&= -\sum_{k=1}^m m\alpha_k^{-1} \int_0^t e^{-\alpha_k(t-x)}(1-x)^{m-1} dx \\
&= -\sum_{k=1}^m m\alpha_k^{-1} e^{-\alpha_k t} \int_0^t e^{\alpha_k x}(1-x)^{m-1} dx \\
&= \sum_{k=1}^m m\alpha_k^{-1} e^{-\alpha_k t}(m-1)! \sum_{j=0}^{m-1} \frac{1 - e^{\alpha_k t}(1-t)^j}{j!} \alpha_k^{j-m} \\
&= m! \sum_{k=1}^m \sum_{j=0}^{m-1} \frac{e^{-\alpha_k t} - (1-t)^j}{j!} \alpha_k^{j-m-1}.
\end{aligned}$$

Now notice that, thanks to Lemma 4.35, we have

$$\sum_{j=0}^{m-1} \sum_{k=1}^m \frac{(1-t)^j}{j!} \alpha_k^{j-m-1} = \sum_{k=1}^m \alpha_k^{-m-1} = \frac{1}{m!}.$$

We also have

$$\sum_{k=1}^m \sum_{j=0}^{m-1} \frac{e^{-\alpha_k t}}{j!} \alpha_k^{j-m-1} = \sum_{k=1}^m \alpha_k^{-m-1} e^{-\alpha_k t} \sum_{j=0}^{m-1} \frac{\alpha^j}{j!} = -\frac{1}{m!} \sum_{k=1}^m \alpha_k^{-1} e^{-\alpha_k t}.$$

The last equation follows from the fact that $\alpha_k$ is a zero of $E_m(x) = \sum_{j=0}^m x^j/j!$. So combining the last two equations with the previous one, we get

$$\mathbb{P}(X_1 \leq t) + m \int_0^t M(t-x)(1-x)^{m-1} dx = -1 - \sum_{k=1}^m \alpha_k^{-1} e^{-\alpha_k t} = M(t). \qquad \square$$

**Corollary 4.36** (Theorem 1.1-(a) in [33]). *The expectation $\mathcal{L}_m$ is given by*

$$\mathcal{L}_m = \sum_{k=1}^m -\alpha_k^{-1} e^{-\alpha_k}.$$

*In particular we have*

$$\mathcal{L}_2 = e(\cos(1) + \sin(1)).$$

*Proof.* Since $L_\infty = 1 + N(1)$, we deduce that $\mathcal{L}_m = 1 + M(1)$ and the result follows immediately from Theorem 4.33. □

**Remark 4.37.** Note that derivatives of $M$ at 0 are the moments of the roots $\alpha_1, \ldots, \alpha_m$ i.e.

$$\mu(j, m) := \sum_{k=1}^{m} \alpha_k^j = (-1)^j M^{(j+1)}(0), \quad \text{for } j \geq 0.$$

The functional equation (4.43) then gives a recursion that these moments satisfy:

$$\mu(-1, m) = 0 \quad \text{and} \quad \mu(j, m) = (m)_{j+1} - \sum_{i=0}^{j-1} (m)_{i+1} \mu(j - i - 1, m), \quad \text{for } j \geq 0.$$

where $(X)_k = X(X-1)\cdots(X-k+1)$ is the $k$-th falling factorial polynomial. These moments are polynomials $\mu(j, \cdot)$ in $m$ and it would be interesting to give an expression for $\mu(j, X)$ and study its properties as suggested in [168].

### 4.5.1 A central limit theorem

For the rest of this section, we are interested in varying the parameter $m$ and studying the behaviour of the distribution of the random variable $L_\infty$ as $m$ grows. To fix some notation, for any integer $m \geq 1$ let $X_1^{(m)}, X_2^{(m)}, \ldots$ be a sequence of i.i.d random variables with beta$(1, m)$ distribution. Let $L_{n,m}$ and $L_{\infty,m}$ denote the following random variables

$$L_{n,m} = \sum_{k=1}^{n} 1\left[S_k^{(m)} \leq 1\right] \quad \text{and} \quad L_{\infty,m} = \sum_{k=1}^{\infty} 1\left[S_k^{(m)} \leq 1\right],$$

with

$$S_n^{(m)} = X_1^{(m)} + \cdots + X_n^{(m)}, \quad \text{for } n \geq 1.$$

The following corollary addresses conjectures 4.1 and 4.2 of [33].

**Corollary 4.38.** *The following central limit theorem holds*

$$\frac{L_{\infty,m} - m}{\sqrt{m}} \to \mathcal{N}(0, 1), \quad \text{as } m \uparrow \infty,$$

*in the topology of weak convergence, where* $\mathcal{N}(0, 1)$ *denotes the normal distribution with mean* 0 *and variance* 1.

*Proof.* For $m \geq 1$ and $x \in \mathbb{R}$ let $u(x,m) := \lfloor m + x\sqrt{m} \rfloor$. We then have

$$\begin{aligned}\mathbb{P}\left(\frac{L_{\infty,m} - m}{\sqrt{m}} \leq x\right) &= \mathbb{P}\left(L_{\infty,m} \leq m + x\sqrt{m}\right) \\ &= \mathbb{P}\left(L_{\infty,m} \leq u(x,m)\right) \\ &= \mathbb{P}\left(S_{u(x,m)+1} > 1\right) \\ &= \mathbb{P}\left(\frac{mS_{u(x,m)+1} - u(x,m)}{\sqrt{u(x,m)}} > \frac{m - u(x,m)}{\sqrt{u(x,m)}}\right).\end{aligned}$$

Now notice that

$$\frac{m - u(x,m)}{\sqrt{u(x,m)}} \to -x \quad \text{as } m \uparrow \infty.$$

and, using the Lindeberg-Feller theorem (see [44, Theorem 3.4.10]) on the triangular array $(Y_{m,k})$ with

$$Y_{m,k} = \frac{1}{\sqrt{m}}(mX_k^{(m)} - 1), \quad \text{for } k \text{ and } m \text{ large enough},$$

we obtain

$$\frac{mS_{u(x,m)+1} - u(x,m)}{\sqrt{u(x,m)}} = \sqrt{\frac{m}{u(x,m)}}\left(Y_{m,1} + \cdots + Y_{m,u(x,m)}\right) \to \mathcal{N}(0,1) \quad \text{as } m \uparrow \infty.$$

So we conclude, as desired, that for any real number $x \in \mathbb{R}$:

$$\mathbb{P}\left(\frac{L_{\infty,m} - m}{\sqrt{m}} \leq x\right) \to \int_{-x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du, \quad \text{as } m \uparrow \infty.$$

Note that the hypothesis of [44, theorem 3.4.10] hold because $mX_1^{(m)}$ converges in distribution to an exponential distribution with parameter 1. □

## 4.6 Wrapping probability distributions on the circle

In the decomposition (4.14) for an exponentially distributed $X = \gamma_1/\lambda$ with parameter $\lambda > 0$; that is

$$\mathbb{P}(X > t) = e^{-\lambda t}, \quad \text{for } t \geq 0,$$

the Eulerian generating function (4.12) is the probability density of the fractional part $(\gamma_1/\lambda)^\circ$ at $u \in [0,1)$. In this probabilistic representation of Euler's exponential generating function (4.3), the factorially normalized Bernoulli polynomials $b_n(u)$ for $n > 0$ are the densities at $u \in [0,1)$ of a sequence of signed measures on $[0,1)$, each with total mass 0, which when weighted by $(-\lambda)^n$ and summed over $n > 0$ give the difference between the probability density of $(\gamma_1/\lambda)^\circ$ and the uniform probability density $b_0(u) \equiv 1$ for $u \in [0,1)$.

For a positive integer $r$ and a positive real number $\lambda$, let $f_{r,\lambda}$ denote the probability density of the gamma$(r,\lambda)$ distribution:

$$f_{\gamma_{r,\lambda}}(x) = \frac{\lambda^r}{\Gamma(r)} e^{-\lambda x} x^{r-1} 1_{x>0}, \quad x \in \mathbb{R}.$$

It is well known that $f_{r,\lambda}$ is the $r$-fold convolution of $f_{1,\lambda}$ on the real line i.e. $f_{\gamma_{r,\lambda}} = (f_{\gamma_{1,\lambda}})^{*r}$. Let $\gamma_{r,\lambda}$ be a random variable with distribution gamma$(r,\lambda)$ and let us denote by $\gamma^\circ_{r,\lambda}$ the random variable $\gamma_{r,\lambda} \mod \mathbb{Z}$ on the circle $\mathbb{T}$. The probability density of $\gamma^\circ_{r,\lambda}$ on $\mathbb{T} = [0,1)$ is given for $0 \le u < 1$ by

$$f^\circ_{\gamma_{r,\lambda}}(u) = \sum_{m\in\mathbb{Z}} f_{\gamma_{r,\lambda}}(u+m) = \frac{\lambda^r}{\Gamma(r)} e^{-\lambda u} \sum_{m=0}^{\infty} (u+m)^{r-1} e^{-\lambda m} \tag{4.46}$$

$$= \frac{\lambda^r}{\Gamma(r)} e^{-\lambda u} \Phi(e^{-\lambda}, 1-r, u), \tag{4.47}$$

where $\Phi$ is the Hurwitz-Lerch zeta function $\Phi(z,s,u) = \sum_{m\ge 0} \frac{z^m}{(u+m)^s}$. In particular, for $r = 1$ the probability density of $\gamma^\circ_{1,\lambda}$, the fractional part of an exponential variable with mean $1/\lambda$, at $u \in [0,1)$, is

$$f^\circ_{\gamma_{1,\lambda}}(u) = \frac{\lambda e^{\lambda(1-u)}}{e^\lambda - 1} = B(1-u,\lambda) = 1 + \sum_{n=1}^{\infty} b_n(1-u)\lambda^n$$

where $B(x,\lambda)$, evaluated here for $x = 1-u$, is the generating function in (4.3). Combined with the reflection symmetry (4.11), this shows that the probability density of $\gamma^\circ_{1,\lambda}$ can be expanded in Bernoulli polynomials as:

$$f^\circ_{\gamma_{1,\lambda}}(u) = 1 + \sum_{n=1}^{\infty} (-1)^n b_n(u)\lambda^n \qquad (0 \le u < 1). \tag{4.48}$$

The following proposition generalizes this result to all integers $r \ge 1$.

The expansion (4.49) can be read from (4.47) and formula (11) on page 30 of [59]. The consequent interpretation (4.50) of $b_r(u)$ for $r > 0$, as the density of a signed measure describing how the probability density $f^\circ_{\gamma_{r,\lambda}}(u)$ approaches the uniform density 1 as $\lambda \downarrow 0$, dates back to the work of Nörlund [130, p. 53], who gave an entirely analytical account of this result. See also [34] for further study of the wrapped gamma and related probability distributions, and [43] for various identities related to (4.49).

**Proposition 4.39** (Wrapped gamma distribution)**.** *For each $r = 1,2,3,\ldots$ the wrapped gamma density admits the following expansion:*

$$f^\circ_{\gamma_{r,\lambda}}(u) = 1 + \sum_{n=r}^{\infty} (-1)^{n-r+1} \binom{n-1}{r-1} b_n(u)\lambda^n \quad \text{for } 0 < \lambda < 2\pi \tag{4.49}$$

*where the convergence is uniform in $u \in [0,1)$. In particular, as $\lambda \downarrow 0$*

$$f^\circ_{\gamma_{r,\lambda}}(u) = 1 - \lambda^r b_r(u) + O(\lambda^{r+1}), \quad \text{uniformly in } u \in [0,1). \tag{4.50}$$

*Proof.* Since $f_{\gamma_{r,\lambda}} = (f_{\gamma_{1,\lambda}})^{*r}$ we deduce that $f^\circ_{\gamma_{r,\lambda}} = (f^\circ_{\gamma_{1,\lambda}})^{\circledast r}$. Then, combining (4.48) and Corollary 4.2 we deduce that

$$\begin{aligned} f^\circ_{\gamma_{r,\lambda}}(u) &= (\underbrace{f^\circ_{\gamma_{1,\lambda}} \circledast \ldots \circledast f^\circ_{\gamma_{1,\lambda}}}_{r \text{ factors}})(u) \\ &= 1 + \sum_{k_1,\ldots,k_r \geq 1} (-1)^{k_1+\cdots+k_r} \lambda^{k_1+\cdots+k_r} (b_{k_1} \circledast \ldots \circledast b_{k_r})(u) \\ &= 1 + \sum_{n=r}^{\infty} \sum_{\substack{k_1,\ldots,k_r \geq 1 \\ k_1+\cdots+k_r=n}} (-1)^n \lambda^n (-1)^{-r+1} b_n(u) \\ &= 1 + \sum_{n=r}^{\infty} (-1)^{n-r+1} A_{r,n} \lambda^n b_n(u), \end{aligned}$$

where $A_{r,n} = \binom{n-1}{r-1}$ is the number of $r$-tuples of positive integers that sum to $n$. Notice that all the sums we considered are summable uniformly in $u \in [0,1]$ since $\|b_n\|_\infty = O((2\pi)^n)$ as $n \to \infty$, see (4.7). □

**Remark 4.40.** The general problem of expanding a function on $\mathbb{T}$ as a sum of Bernoulli polynomials was first treated by Jordan [95, Section 85] and Mordell [126]. In our context, we think of the expansion of a function in Bernoulli polynomials as an analog of the Taylor expansion where we work with the convolutions $\circledast$ instead of the usual multiplication of functions; i.e. we view expansions of the form

$$f(x) = a_0(f) + \sum_{n=1}^{\infty} (-1)^{n-1} a_n(f) b_1^{\circledast n}(x) = a_0(f) + \sum_{n=1}^{\infty} a_n(f) b_n(x),$$

as an analogue of Taylor expansions

$$f(x) = f(0) + \sum_{n=1}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

As we have seen in this section, this point of view is especially fruitful when one wishes to convolve probability measures on $\mathbb{T} = [0,1)$. If $f$ is a $C^\infty$ function on $[0,1]$ satisfying some dominance condition (see [126, Theorem 1]), the coefficient of $b_1^{\circledast}(x)$ in the expansion of $f$ is given by

$$(-1)^{n-1} a_n(f) = (f^{(n-1)}(1) - f^{(n-1)}(0)), \quad \text{for } n \geq 0.$$

## 4.7 Conclusion

To conclude, this chapter underlines a property of Bernoulli polynomials in terms of circular convolution which, surprisingly, was not recorded in the vast literature on the topic.

There is a probabilistic and combinatorial model underlying this property for which we coined the name Bernoulli clock. With this model in mind, this chapter offers a pleasant probabilistic perspective to the work [91] of Horton and Kurn and the recent work of Clifton et al. [33] on counting permutations of the mutliset $1^m \cdots n^m$.

# Part II

# Number theory, Combinatorics and Geometry

# Chapter 5

# Orders and convex sets in Bruhat-Tits Buildings

This chapter is based joint work [52, 55] with Marvin A. Hahn, Gabriele Nebe, Mima Stanojkovski and Bernd Sturmfels. The content of this chapter appeared (in modified form) in the journal *Beiträge zur Algebra und Geometrie* (for [55]) and in the *International Journal of Number Theory* (for [52]).

In this chapter, we study orders in the ring of $d \times d$ matrices $K^{d\times d}$ over a discretely valued field $K$. We shall also discuss their action on the Bruhat-Tits building $\mathcal{B}_d(K)$ and describe the set of fixed points under this action.

## 5.1 Introduction

Throughout this chapter, let $K$ be a non-archimedean valued field with a surjective discrete valuation $\mathrm{val} : K \to \mathbb{Z} \cup \{\infty\}$, with valuation ring $\mathcal{O}_K$, unifomizer $\varpi$, and maximal ideal $\mathfrak{m}_K = \varpi\mathcal{O}_K$. There is no need for $K$ to be complete: in particular, $K = \mathbb{Q}$ with some $p$-adic valuation is allowed. We refer the reader to Chapter 1 for a brief introduction to non-archimedean valued fields. Finally, we fix a positive integer $d$.

An *order* $\Lambda$ in $K^{d\times d}$ is a finitely generated $\mathcal{O}_K$-submodule of $K^{d\times d}$ that is also a subring of $K^{d\times d}$ (that is $\Lambda$ contains the identity matrix $I_d$ and is closed under matrix multiplication). An order $\Lambda$ is *maximal* if it is not properly contained in any other order. One example of a maximal order is the matrix ring

$$\mathcal{O}_K^{d\times d} := \{X \in K^{d\times d} : \mathrm{val}(x_{ij}) \geq 0 \text{ for all } 1 \leq i, j \leq d\}.$$

If $\Lambda$ is an order in $K^{d\times d}$ and $L \subset K^d$ is a lattice, we define $\Lambda \cdot L$ as follows:

$$\Lambda \cdot L = \{g \cdot x \colon g \in \Lambda \text{ and } x \in L\}.$$

This defines an action of $\Lambda$ on points of the Bruhat-Tits building $\mathcal{B}_d(K)$. See Section 1.2 and Section 2.5.

**Definition 5.1.** Let $\Lambda$ be an order in $K^{d\times d}$. Then

$$Q(\Lambda) := \{[L] \in \mathcal{B}_d^0(K) \colon \Lambda \cdot L = L\}$$

denotes the set of homothety classes of $\Lambda$-lattices in $K^d$. The order $\Lambda$ is called *closed* if

$$\Lambda = \bigcap_{[L]\in Q(\Lambda)} \mathrm{End}_{\mathcal{O}_K}(L).$$

Notice that the closed orders are exactly the ones that are determined by their sets of invariant lattices. Not all order are closed as we shall see in Section 5.3 (see Remark 5.45 for an example).

**Definition 5.2.** Given a finite collection $\Gamma = \{[L_1], \dots, [L_s]\}$ of lattice classes in the building $\mathcal{B}_d(K)$, the *Plesken-Zassenhaus order* of $\Gamma$ is the order defined by

$$\mathrm{PZ}(\Gamma) = \bigcap_{i=1}^{s} \mathrm{End}_{\mathcal{O}_K}(L_i).$$

In this chapter we are interested in studying orders in $K^{d\times d}$ and their action on the Bruhat-Tits building $\mathcal{B}_d(K)$. In particular, given an order $\Lambda$, we are interested in describing the set of lattice classes that are invariant under the action of $\Lambda$. Conversely, given a finite collection $\Gamma = \{[L_1], \dots, [L_s]\}$ of lattice classes in the building $\mathcal{B}_d(K)$, we study is associated Plesken-Zassenhaus order $\mathrm{PZ}(\Gamma)$.

This chapter is organized as follows. In Section 5.2, we study the case where the configuration $\Gamma$ lies in a common apartment $\mathcal{A}$ of the building $\mathcal{B}_d(K)$. Here, the associated order $\Lambda = \mathrm{PZ}(\Gamma)$ is called a *graduated order* following [133]. The set of fixed points in $\mathcal{B}_d$ under the action of $\Lambda$ is *tropically convex* in the apartment $\mathcal{A}$, i.e. a *polytrope* in the *tropical torus* $\mathbb{R}^d/\mathbb{R}$, see [96, Section 6.5]. Subsection 5.2.1 concerns graduated orders in $K^{d\times d}$. In Proposition 5.8 and Proposition 5.10 we present linear inequalities that characterize these orders and the lattices they act on. These inequalities play an important role in tropical convexity (see [96, Chapter 5]), to be explained in Subsection 5.2.2. Theorem 5.14 gives a tropical matrix formula for the Plesken-Zassenhaus order of a collection of diagonal lattices. In Subsection 5.2.3 we introduce polytrope regions. These are convex cones and polyhedra whose integer points represent graduated orders. Subsection 5.2.4 is concerned with (fractional) ideals in an order $\Lambda_M$. These are parametrized by the ideal class polytrope $\mathcal{Q}_M$. In Subsection 5.2.5 we turn to Bruhat-Tits buildings and their chambers.

In Section 5.3, the second part of this chapter, we extend our study to configurations of lattices $\Gamma$ in the building $\mathcal{B}_d(K)$ that are given by the *Minkowski sum* of a polytrope (in some apartment $\mathcal{A}$) and a ball of radius $r$ in $\mathcal{B}_d(K)$. We call the orders associated to such a configuration *bolytrope orders*. We introduce a notion of distance in the Bruhat-Tits building in Subsection 5.3.1. In Subsection 5.3.2 we define bolytropes and present our main tool: the radical idealizer chain of an order in Subsection 5.3.3. Subsection 5.3.4 is dedicated to ball orders and bolytrope orders. Finally, in Subsection 5.3.6 we discuss the case when $d = 2$ in which the building is an infinite tree.

## 5.2 Graduated orders

We write $K^{d\times d}$ for the ring of $d \times d$ matrices with entries in $K$. The map val is applied coordinatewise to matrices and vectors. For example, if $K = \mathbb{Q}$ with $p = 2$, then the vector $x = (8/7, 5/12, 17)$ has $\mathrm{val}(x) = (3, -2, 0)$. In what follows, we often take $X = (x_{ij})$ to be a $d \times d$ matrix with nonzero entries in $K$. In this case, $\mathrm{val}(X) = (\mathrm{val}(x_{ij}))$ is a matrix in $\mathbb{Z}^{d\times d}$.

Fix any square matrix $M = (m_{ij})$ in $\mathbb{Z}^{d\times d}$. This section revolves around the interplay between the following two objects associated with $M$, one algebraic and the other geometric:

1. the set $\Lambda_M = \{X \in K^{d\times d} : \mathrm{val}(X) \geq M\}$, an $\mathcal{O}_K$-lattice in the vector space $K^{d\times d}$;
2. the set $Q_M = \{u \in \mathbb{R}^d/\mathbb{R}\mathbf{1} : u_i - u_j \leq m_{ij} \text{ for } 1 \leq i, j \leq d\}$, where $\mathbf{1} = (1, \ldots, 1)$.

This interplay is strongest and most interesting when $\Lambda_M$ is closed under multiplication. In this case, $\Lambda_M$ is a non-commutative ring of matrices. Such a ring is called an order in $K^{d\times d}$. The quotient space $\mathbb{R}^d/\mathbb{R}\mathbf{1} \simeq \mathbb{R}^{d-1}$ is the usual setting for tropical geometry [96, 116]. Note that $Q_M$ is a convex polytope in that space. It is also tropically convex, for both the min-plus algebra and the max-plus algebra. Following [96, 161], we use the term *polytrope* for $Q_M$.

**Example 5.3.** For $d = 4$, fix the matrix with diagonal entries 0 and off-diagonal entries 1:

$$M = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}. \tag{5.1}$$

The polytrope $Q_M$ is the set of solutions to the 12 inequalities $u_i - u_j \leq 1$ for $i \neq j$. It is the 3-dimensional polytope shown in Figure 5.1. Namely, $Q_M$ is a *rhombic dodecahedron*, with 14 vertices, 24 edges and 12 facets. The vertices are the images in $\mathbb{R}^4/\mathbb{R}\mathbf{1}$ of the 14 vectors in $\{0, 1\}^4\backslash\{\mathbf{0}, \mathbf{1}\}$. Vertices $e_i$ are blue, vertices $e_i + e_j$ are yellow, and vertices $e_i + e_j + e_k$ are red.

The order $\Lambda_M$ consists of all $4 \times 4$ matrices with entries in the valuation ring $\mathcal{O}_K$ whose off-diagonal elements lie in the maximal ideal $\mathfrak{m}_K = \varpi\mathcal{O}_K$. We shall see in Theorem 5.20 that the blue and red vertices encode the injective modules and the projective modules of $\Lambda_M$ respectively.

The connection between algebra, geometry and combinatorics we present was pioneered by Plesken and Zassenhaus. Our primary source on their work is the book [133]. One objective of this section is to give an exposition of their results using the framework of tropical geometry [96, 116]. But we also present a range of new results.

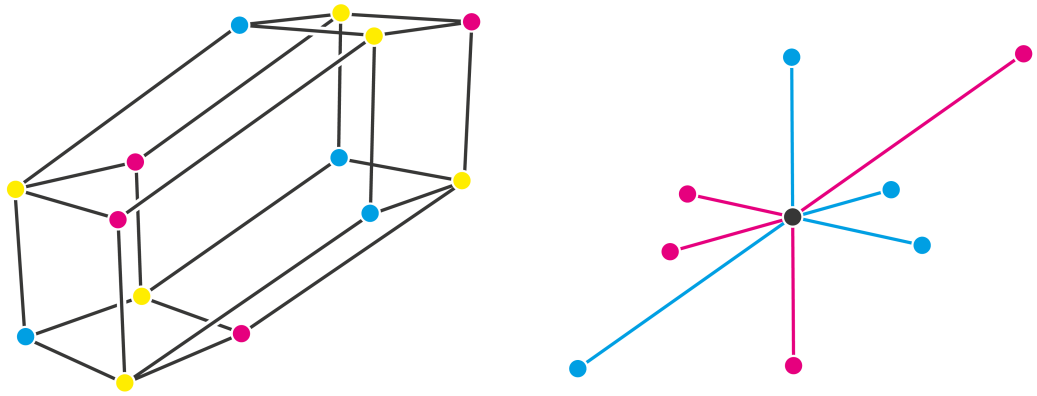


Figure 5.1: The polytrope $Q_M$ on the left is a rhombic dodecahedron. The four blue vertices and the four red vertices, highlighted on the right, will play a special role for the order $\Lambda_M$.

### 5.2.1 Graduated orders

We begin with some standard facts found that can be found in [133]. The first is a natural bijection between lattice classes $[L]$ in $K^d$ and maximal orders in $K^{d\times d}$.

**Proposition 5.4.** *Any order $\Lambda$ in $K^{d\times d}$ is contained in the endomorphism ring of a lattice $L \subset K^d$. The maximal orders in $K^{d\times d}$ are exactly the endomorphism rings of lattices $L$:*

$$\mathrm{End}_{\mathcal{O}_K}(L) \;:=\; \{X \in K^{d\times d} : XL \subseteq L\}.$$

*Two lattices $L$ and $L'$ in $K^d$ are equivalent if and only if $\mathrm{End}_{\mathcal{O}_K}(L) = \mathrm{End}_{\mathcal{O}_K}(L')$.*

*Proof.* Let $\Lambda = \bigoplus_{j=1}^{d^2} \mathcal{O}_K X_j$ be an order in $K^{d\times d}$. If we apply all the matrices $X_j$ to the *standard lattice* $L_0 = \mathcal{O}_K^d = \bigoplus_{i=1}^{d} \mathcal{O}_K e_i$, then we obtain the following lattice in $K^d$:

$$L \;:=\; \sum_{j=1}^{d^2} X_j L_0 \;=\; \sum_{i=1}^{d}\sum_{j=1}^{d^2} \mathcal{O}_K X_j\, e_i.$$

Since $\Lambda$ is closed under multiplication, we have $X_j L \subseteq L$ for all $j$. Therefore $\Lambda \subseteq \mathrm{End}_{\mathcal{O}_K}(L)$.

Endomorphism rings of lattices are orders. Indeed, if $L = gL_0$ for $g \in \mathrm{GL}(d, K)$, then

$$\mathrm{End}_{\mathcal{O}_K}(L) = g\,\mathrm{End}_{\mathcal{O}_K}(L_0)g^{-1} = g\;\mathcal{O}_K^{d\times d}\;g^{-1}. \tag{5.2}$$

This is a ring, and it is spanned as an $\mathcal{O}_K$-lattice by $\{gE_{ij}g^{-1} : 1 \le i, j \le d\}$. This allows to conclude that the maximal orders are exactly the endomorphism rings of lattices. □

For $u \in \mathbb{Z}^d$ we abbreviate $g_u = \mathrm{diag}(\varpi^{u_1}, \varpi^{u_2}, \cdots, \varpi^{u_d})$. This diagonal matrix transforms the standard lattice $\mathcal{O}_K^d$ to $L_u = g_u\mathcal{O}_K^d$. The endomorphism ring $\mathrm{End}_{\mathcal{O}_K}(L_u)$ is the maximal order in (5.2). Let $M(u)$ denote the $d \times d$ matrix whose entry in position $(i, j)$ equals $u_i - u_j$.

**Lemma 5.5.** *The endomorphism ring of the lattice $L_u$ is given by valuation inequalities:*

$$\mathrm{End}_{\mathcal{O}_K}(L_u) \;=\; \Lambda_{M(u)} \;=\; \{\, X \in K^{d\times d} \,:\, \mathrm{val}(X) \geq M(u) \,\}. \tag{5.3}$$

*Proof.* The elements of $\mathrm{End}_{\mathcal{O}_K}(L_u)$ are the matrices $X = g_u Y g_u^{-1}$ where $Y \in \mathcal{O}_K^{d\times d}$. Writing $X = (x_{ij})$ and $Y = (y_{ij})$, the equation $X = g_u Y g_u^{-1}$ means that $x_{ij} = \varpi^{u_i - u_j} y_{ij}$ for all $i, j$. The condition $\mathrm{val}(y_{ij}) \geq 0$ is equivalent to $\mathrm{val}(x_{ij}) \geq u_i - u_j$. Taking the conjunction over all $(i, j)$, we conclude that $\mathrm{val}(Y) \geq 0$ is equivalent to the desired inequality $\mathrm{val}(X) \geq M(u)$. □

The matrices $M(u)$ are characterized by the following two properties. All diagonal entries are zero and the tropical rank is one, cf. [116, Section 5.3]. What happens if we replace $M(u)$ in (5.3) by an arbitrary matrix $M \in \mathbb{Z}^{d\times d}$? Then we get the set $\Lambda_M$ from the begining of Section 5.2.

**Remark 5.6.** For any matrix $M \in \mathbb{Z}^{d\times d}$, the set $\Lambda_M$ is a lattice in $K^{d\times d}$. It is generated as an $\mathcal{O}_K$-module by the matrices $\varpi^{m_{ij}} E_{ij}$ for $1 \leq i, j \leq d$. The lattice $\Lambda_M$ may not be an order.

Write $\mathbb{Z}_0^{d\times d}$ for the set of integer matrices $M$ with zeros on the diagonal, i.e. $m_{ii} = 0$ for all $i$. If $M$ lies in $\mathbb{Z}_0^{d\times d}$ then $\Lambda_M$ contains the identity matrix, but may still not be an order.

**Example 5.7.** Let $K = \mathbb{Q}$ with the $p$-adic valuation, for some prime $p \geq 5$. For $d = 3$, set

$$M = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 1 & 1 & p \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \text{so} \quad X^2 = \begin{bmatrix} 2+p & 2+p & 1+2p \\ 3 & 3 & 2+p \\ 3 & 3 & 2+p \end{bmatrix}.$$

Since $\mathrm{val}(X) = M$ and $\mathrm{val}(X^2) = 0$, we have $X \in \Lambda_M$ but $X^2 \notin \Lambda_M$. This means that $\Lambda_M$ is not an order.

**Proposition 5.8.** *Given $M = (m_{ij})$ in $\mathbb{Z}_0^{d\times d}$, the lattice $\Lambda_M$ is an order in $K^{d\times d}$ if and only if $M$ satisfies the following inequalities*

$$m_{ij} + m_{jk} \geq m_{ik} \quad \textit{for all } 1 \leq i, j, k \leq d. \tag{5.4}$$

*Proof.* To prove the if direction, we assume (5.4). Our hypothesis $m_{ii} = 0$ ensures that $\Lambda_M$ contains the identity matrix, so $\Lambda_M$ has a multiplicative unit. Suppose $X, Y \in \Lambda_M$. Then the $(i, k)$ entry of $XY$ equals $\sum_{j=1}^{d} x_{ij} y_{jk}$. This is a scalar in $K$ whose valuation is at least $m_{ij} + m_{jk}$ for some index $j$. Hence it is greater than or equal to $m_{ik}$ since (5.4) holds.

For the only-if direction, suppose $m_{ij} + m_{jk} < m_{ik}$. Then $X = \varpi^{m_{ij}} E_{ij}$ and $Y = \varpi^{m_{jk}} E_{jk}$ are in $\Lambda_M$. However, $XY = \varpi^{m_{ij}+m_{jk}} E_{ik}$ is not in $\Lambda_M$ because its entry in position $(i, k)$ has valuation less than $m_{ik}$. Hence $\Lambda_M$ is not multiplicatively closed, so it is not an order. □

In what follows, we define graduated orders following [133] and collect some related results from [55, 133].

**Definition 5.9.** An $\mathcal{O}_K$-order $\Lambda$ in $K^{d\times d}$ is called *graduated* if $\Lambda$ contains a complete set of orthogonal primitive idempotents $\epsilon_1, \dots, \epsilon_d$ of $K^{d\times d}$.

The primitive idempotents of $K^{d\times d}$ are exactly the projections onto 1-dimensional subspaces of $K^d$, so each set $\{\epsilon_1, \dots, \epsilon_d\}$ as in Definition 5.9 defines a *frame*

$$K^d = \epsilon_1 K^d \oplus \dots \oplus \epsilon_d K^d = Ke_1 \oplus \dots \oplus Ke_d$$

i.e. a decomposition of $K^d$ as a direct sum of 1-dimensional subspaces. In any frame basis $(e_1, \dots, e_d)$ the idempotents are diagonal matrices with exactly one entry 1 on the diagonal. The projection onto the $ij$-matrix entry $\epsilon_i \Lambda \epsilon_j$ is an $\mathcal{O}_K$-submodule of $\epsilon_i K^{d\times d} \epsilon_j \cong K$. Hence, writing matrices with respect to the frame basis $(e_1, \dots, e_d)$, there exists a matrix $M = (m_{ij}) \in \mathbb{Z}^{d\times d}$ such that the graduated order $\Lambda$ is of the form

$$\Lambda_M = \{X = (X_{ij}) \in K^{d\times d} \colon X_{ij} \in \mathfrak{m}_K^{m_{ij}} \text{ for all } i, j = 1, \dots, d\}.$$

The matrix $M$ is called the *exponent matrix* of $\Lambda$.

Fix $M$ that satisfies (5.4). The graduated order $\Lambda_M$ is an $\mathcal{O}_K$-subalgebra of $K^{d\times d}$. It is therefore natural to ask which lattices in $K^d$ are $\Lambda_M$-stable.

**Proposition 5.10.** *A lattice $L$ is stable under $\Lambda_M$ if and only if $L = L_u$ with $u \in \mathbb{Z}^d$ satisfying*

$$u_i - u_j \le m_{ij} \quad \text{for} \quad 1 \le i, j \le d. \tag{5.5}$$

*Moreover, if $u, u' \in \mathbb{Z}^d$ satisfy (5.5), then the diagonal lattices $L_u$ and $L_{u'}$ are isomorphic as $\Lambda_M$-modules if and only if they are equivalent, i.e. $u = u'$ in the quotient space $\mathbb{R}^d/\mathbb{R}\mathbf{1}$.*

*Proof.* Fix a lattice $L$ and let $u = (u_1, \dots, u_d)$ be defined by $u_i = \min\{\mathrm{val}(b_i) : b \in L\}$. Then $L \subseteq L_u$ because every $b \in L$ is an $\mathcal{O}_K$-linear combination of the standard basis of $L_u$, namely $b = \sum_{i=1}^d b_i e_i = \sum_{i=1}^d (b_i \varpi^{-u_i}) \varpi^{u_i} e_i$. Suppose that $L$ is $\Lambda_M$-stable. Since $m_{ii} = 0$, we have $E_{ii} \in \Lambda_M$. Hence $E_{ii}\, b = b_i e_i \in L$ for every $b \in L$. This implies $L_u \subseteq L$ and hence $L = L_u$. Applying $\varpi^{m_{ij}} E_{ij} \in \Lambda_M$ to $\varpi^{u_j} e_j \in L_u$, we see that $\varpi^{m_{ij}+u_j} e_i$ lies in $L_u$, and this implies $m_{ij} + u_j \ge u_i$. Hence (5.5) holds. Conversely, suppose that (5.5) holds. Then the generator $\varpi^{m_{ij}} E_{ij}$ of $\Lambda_M$ maps each basis vector $\varpi^{u_k} e_k$ of $L_u$ either to zero (if $j \ne k$), or to $\varpi^{m_{ik}+u_k} e_i \in L_u$. This proves the first assertion.

For the second assertion, let $u, u' \in \mathbb{Z}^d$ satisfy (5.5). Since multiplication by $\alpha \in K^*$ is an isomorphism of $\mathcal{O}_K$-modules, the if-direction is clear. Conversely, if $L_u$ and $L_{u'}$ are isomorphic, then there exists $g \in \mathrm{GL}_d(K)$ such that $L_{u'} = gL_u$ and $gX = Xg$ for all $X \in \Lambda_M$. Pick $s \in \mathbb{Z}_{>0}$ such that $\varpi^s \mathcal{O}_K^{d\times d} \subset \Lambda_M$. Then $g$ commutes with every matrix in $\varpi^s \mathcal{O}_K^{d\times d}$. This implies that $g$ is central in $\mathcal{O}_K^{d\times d}$, and therefore $g$ is a multiple of the identity matrix. □

**Remark 5.11.** For $M = (m_{ij}) \in \mathbb{Z}_0^{d\times d}$ satisfying (5.4), the $\Lambda_M$-lattices $L$ are of the form $L = \bigoplus_{i=1}^d \epsilon_i L$ and hence there exists $u = (u_1, \dots, u_d) \in \mathbb{Z}^d$ such that

$$L = L_u := \mathcal{O}_K \varpi^{u_1} e_1 \oplus \dots \oplus \mathcal{O}_K \varpi^{u_d} e_d.$$

The tuple $u$ is called the *exponent vector* of the lattice $L$. Moreover, $L = L_u$ is a $\Lambda_M$-lattice if and only if, for any choice of $1 \leq i, j \leq d$, one has $m_{ij} + u_j \geq u_i$ and two $\Lambda_M$-lattices $L_u$ and $L_v$ are isomorphic if and only if $u - v \in \mathbb{Z}\mathbf{1}$.

In the next section we shall see that $\Lambda_M$-lattices are parameterized by the integral points of the set

$$Q_M := \{[u] \in \mathbb{R}^d/\mathbb{R}\mathbf{1} \colon m_{ij} + u_j \geq u_i \text{ for all } i, j = 1, \ldots, d\}.$$

Following [96], the set $Q_M$ is called a *polytrope*.

### 5.2.2 Bi-tropical Convexity

We now develop the relationship between graduated orders and tropical mathematics [96, 116]. Both the *min-plus algebra* $(\mathbb{R}, \underline{\oplus}, \odot)$ and the *max-plus algebra* $(\mathbb{R}, \overline{\oplus}, \odot)$ will be used. Its arithmetic operations are the minimum, maximum, and classical addition of real numbers:

$$a \,\underline{\oplus}\, b \,=\, \min(a,b)\,, \;\; a \,\overline{\oplus}\, b \,=\, \max(a,b)\,, \;\; a \odot b \,=\, a+b \qquad \text{for} \qquad a, b \in \mathbb{R}.$$

If $M$ and $N$ are real matrices, and the number of columns of $M$ equals the number of rows of $N$, then we write $M \,\underline{\odot}\, N$ and $M \,\overline{\odot}\, N$ for their respective matrix products in these algebras.

**Example 5.12.** Consider the $2 \times 2$ matrices $M = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$ and $N = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. We find that

$$M \,\underline{\odot}\, M = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \quad M \,\underline{\odot}\, N = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad N \,\underline{\odot}\, M = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad N \,\underline{\odot}\, N = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$
$$M \,\overline{\odot}\, M = \begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix}, \quad M \,\overline{\odot}\, N = \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix}, \quad N \,\overline{\odot}\, M = \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}, \quad N \,\overline{\odot}\, N = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}.$$

There are two flavors of tropical convexity [116, Section 5.2]. A subset of $\mathbb{R}^d$ is *min-convex* if it is closed under linear combinations in the min-plus algebra, and *max-convex* if the same holds for the max-plus algebra. Thus convex sets are images of matrices under linear maps.

We are especially interested in bi-tropical convexity in the ambient space $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. This is ubiquitous in [96, Section 5.4] and [116]. Joswig [96, Section 1.4] calls it the *tropical projective torus*. At a later stage, we also work in the corresponding matrix space $\mathbb{R}^{d\times d}/\mathbb{R}\mathbf{1}$.

Let $\mathbb{R}_0^{d\times d}$ denote the space of real $d \times d$ matrices with zeros on the diagonal, which is a real $(d^2 - d)$-dimensional vector space with lattice $\mathbb{Z}_0^{d\times d}$. For $M = (m_{ij})$ in $\mathbb{R}_0^{d\times d}$, we define

$$Q_M \;=\; \left\{ u \in \mathbb{R}^d/\mathbb{R}\mathbf{1} \,:\, u_i - u_j \;\leq\; m_{ij} \text{ for } 1 \leq i, j \leq d \right\}. \tag{5.6}$$

Such a set is known as a *polytrope* in tropical geometry [97, 116]. Other communities use the terms *alcoved polytope* and *weighted digraph polytope*. We note that $Q_M$ is both min-convex and max-convex [96, Proposition 5.30] and, being a polytope, it is also classically convex.

Using tropical arithmetic, the linear inequalities in (5.4) can be written concisely as follows:

$$M \odot M \;=\; M. \tag{5.7}$$

Thus, $M$ is *min-plus idempotent*. This holds for $M$ in Example 5.12. Joswig's book [96, Section 3.3] uses the term *Kleene star* for matrices $M \in \mathbb{R}_0^{d\times d}$ with (5.7). Propositions 5.8 and 5.10 imply:

**Corollary 5.13.** *The lattice $\Lambda_M$ is an order in $K^{d\times d}$ if and only if (5.7) holds. In this case, the integer points $u$ in the polytrope $Q_M$ are in bijection with the isomorphism classes of $\Lambda_M$-lattices $L_u$. Here, by a $\Lambda_M$-*lattice *we mean a $\Lambda_M$-module that is also a lattice in $K^d$.*

Let $\Gamma = \{L_1, \ldots, L_n\}$ be a finite set of lattices in $K^d$, which might be taken up to equivalence. The intersection of two orders in $K^{d\times d}$ is again an order. Hence the intersection

$$\mathrm{PZ}(\Gamma) \;=\; \mathrm{End}_{\mathcal{O}_K}(L_1) \cap \cdots \cap \mathrm{End}_{\mathcal{O}_K}(L_n) \tag{5.8}$$

is an order in $K^{d\times d}$. We call $\mathrm{PZ}(\Gamma)$ the *Plesken-Zassenhaus order* of the configuration $\Gamma$.

In the following we assume that each $L_i$ is a *diagonal lattice*, i.e. $L_i = L_{u^{(i)}}$ for $u^{(i)} \in \mathbb{Z}^d$. Our next result involves a curious mix of max-plus algebra and min-plus algebra.

**Theorem 5.14.** *Let $\Gamma = \{L_{u^{(1)}}, \ldots, L_{u^{(n)}}\}$ be any configuration of diagonal lattices in $K^d$. Then its Plesken-Zassenhaus order $PZ(\Gamma)$ coincides with the graduated order $\Lambda_M$ where*

$$M \;=\; M(u^{(1)}) \,\overline{\oplus}\, M(u^{(2)}) \,\overline{\oplus}\, \cdots \,\overline{\oplus}\, M(u^{(n)}). \tag{5.9}$$

*This max-plus sum of tropical rank one matrices is min-plus idempotent, i.e. (5.4) and (5.7) hold.*

*Proof.* We regard $\Gamma$ as a configuration in $\mathbb{R}^d/\mathbb{R}\mathbf{1}$. By construction, $M$ is the entrywise smallest matrix such that $\Gamma$ is contained in the polytrope $Q_M$. From [96, Lemma 3.25] the matrix $M$ is a Kleene star, that is (5.4) and (5.7) hold. The intersection in (5.8) is defined by the conjunction of the $n$ inequalities $\mathrm{val}(X) \geq M(u^{(i)})$, which is equivalent to $\mathrm{val}(X) \geq M$. □

**Example 5.15.** For $d = n = 3$, fix $u^{(1)} = (-2,-1,0)$, $u^{(2)} = (2,1,0)$, $u^{(3)} = (-1,3,0)$ in $\mathbb{R}^3/\mathbb{R}\mathbf{1}$. The configuration $\Gamma = \{u^{(1)}, u^{(2)}, u^{(3)}\}$ consists of the three red points in Figure 5.2. The red diagram is their min-plus convex hull. This tropical triangle consists of a classical triangle together with three red line segments connected to $\Gamma$. This red min-plus triangle is not convex. The green shaded hexagon is the polytrope spanned by $\Gamma$. By [96, Remark 5.33], this is the geodesic convex hull of $\Gamma$. It equals $Q_M$ where $M$ is computed by (5.9) as follows:

$$M \;=\; (u^{(1)})^t \odot (-u^{(1)}) \;\overline{\oplus}\; (u^{(2)})^t \odot (-u^{(2)}) \;\overline{\oplus}\; (u^{(3)})^t \odot (-u^{(3)}) \;=\; \begin{bmatrix} 0 & 1 & 2 \\ 4 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix}.$$

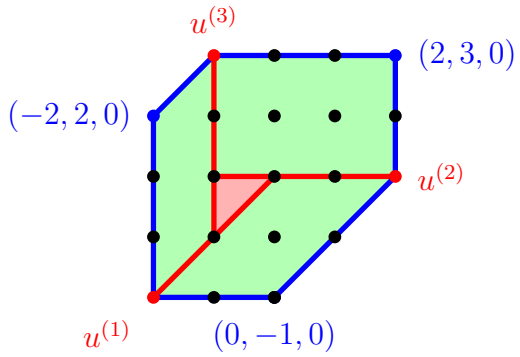



Figure 5.2: A polytrope with three min-plus vertices (blue) and three max-plus vertices (red).

The polytrope $Q_M$ is both a min-plus triangle and a max-plus triangle. Its min-plus vertices, shown in blue, are equal in $\mathbb{R}^3/\mathbb{R}\mathbf{1}$ to the columns of $M$. Its max-plus vertices, shown in red, are the points $u^{(i)}$. These are equal in $\mathbb{R}^3/\mathbb{R}\mathbf{1}$ to the columns of $-M^t$. See Theorem 5.20.

**Remark 5.16.** All lattices $L_u$ for $u \in Q_M$ are indecomposable as $\Lambda_M$-modules, cf. [133]. This is no longer true if $\mathbb{R}$ is enlarged to the tropical numbers $\mathbb{R} \cup \{\infty\}$. The combinatorial theory of polytropes in [96] is set up for this extension, and it indeed makes sense to study orders $\Lambda_M$ with $m_{ij} = \infty$. While we do not pursue this here, our approach would extend to that setting.

**Example 5.17.** Set $d = 4$. The rhombic dodecahedron in Example 5.3 was called the *pyrope* in [97, Figure 4]. This $Q_M$ is a tropical tetrahedron for both min-convexity and max-convexity. The respective vertices are shown in red and blue in Figure 5.1. We have $\Lambda_M = \mathrm{PZ}(\Gamma)$ where $\Gamma$ is either set of four vertices. The $\Lambda_M$-lattices $L_u$ correspond to the 15 integer points in $Q_M$.

### 5.2.3 Polytrope Regions

We next introduce a cone that parametrizes all graduated orders $\Lambda_M$. Following Tran [161], the *polytrope region* $\mathcal{P}_d$ is the set of all min-plus idempotent matrices $M \in \mathbb{R}_0^{d\times d}$. Thus, $\mathcal{P}_d$ is the $(d^2-d)$-dimensional convex polyhedral cone defined by the linear inequalities in (5.4). The equations $m_{ik} = m_{ij} + m_{jk}$ define the cycle space of the complete bidirected graph $\mathcal{K}_d$. This is the lineality space of $\mathcal{P}_d$. Modulo this $(d-1)$-dimensional space, the polytrope region $\mathcal{P}_d$ is a pointed cone of dimension $(d-1)^2$. We view it as a polytope of dimension $d^2-2d$. Each inequality $m_{ik} \leq m_{ij}+m_{jk}$ is facet-defining, so the number of facets of $\mathcal{P}_d$ is $d(d-1)(d-2)$.

It is interesting but difficult to list the vertices of $\mathcal{P}_d$ and to explore the face lattice. The same problem was studied in [12] for the *metric cone*, which is the restriction of $\mathcal{P}_d$ to

the subspace of symmetric matrices in $\mathbb{R}_0^{d\times d}$. A website maintained by Antoine Deza [38] reports that the number of rays of the metric cone equals $3, 7, 25, 296, 55226, 119269588$ for $d = 3, 4, 5, 6, 7, 8$. We here initiate the census for the polytrope region. The following tables report the size of the orbit, the number of incident facets, and a representative matrix $[m_{ij}]$. Here orbit and representatives refer to the natural action of the symmetric group $S_d$ on $\mathcal{P}_d$. The matrices $[m_{ij}]$ in $\mathbb{Z}_0^{3\times 3}$ are written in the vectorized format $[m_{12}m_{13}m_{21}m_{23}m_{31}m_{32}]$.

**Proposition 5.18.** *The polytope $\mathcal{P}_3$ is a bypramid, with f-vector $(5, 9, 6)$. Its five vertices are*

$$3\,,\,4\quad[001100]\quad\text{and}\quad 2\,,\,3\quad[001110].$$

*The polytope $\mathcal{P}_4$ has the f-vector $(37, 327, 1140, 1902, 1680, 808, 204, 24)$. Its 37 vertices are*

| | | | | | |
|---|---|---|---|---|---|
| $12, 10$ | [111011001001] | $6, 12$ | [111011001000] | $12, 14$ | [011011001000] |
| $3, 16$ | [011011000000] | $4, 18$ | [111000000000]. | | |

*The corresponding polytropes $Q_M$ are pyramid, tetrahedron, triangle, segment, and segment. The 15-dimensional polytope $\mathcal{P}_5$ has 2333 vertices in 33 symmetry classes. These classes are*

| | | | | | |
|---|---|---|---|---|---|
| $5, 48$ | [000000000000000001111] | $10, 18$ | [000010012111211111100] | $10, 42$ | [000000000000111011110] |
| $20, 15$ | [000020123232310122 01] | $20, 21$ | [000010001100211121 11] | $20, 39$ | [000000000000111011111] |
| $24, 20$ | [000010012101221111 10] | $24, 30$ | [000010001100111011 11] | $30, 24$ | [000010001101211111 10] |
| $30, 30$ | [000000001100111111 11] | $30, 30$ | [000000001101111111 10] | $30, 36$ | [000000001100110011 11] |
| $40, 18$ | [000020002212222122 12] | $60, 18$ | [000010002101221121 10] | $60, 18$ | [000010012101221211 00] |
| $60, 22$ | [000010001101221111 10] | $60, 27$ | [000010001100111021 11] | $60, 29$ | [000000001100111022 11] |
| $60, 33$ | [000000001100111011 11] | $120, 16$ | [000010012201321221 10] | $120, 17$ | [000010012101221221 10] |
| $120, 18$ | [000010012101221121 10] | $120, 18$ | [000010012101221222 10] | $120, 18$ | [000010012102221221 10] |
| $120, 18$ | [000010012201322132 10] | $120, 19$ | [000010002100221032 21] | $120, 19$ | [000010002101221221 10] |
| $120, 19$ | [000010012101222122 10] | $120, 22$ | [000010001100211022 21] | $120, 22$ | [000010001101221211 10] |
| $120, 23$ | [000010001100211022 11] | $120, 23$ | [000010001100211022 22] | $120, 25$ | [000010001100111022 11] |

*Proof.* This was found by computations with `Maple` and `Polymake` [11]. □

**Remark 5.19.** The integer matrices $M$ in the polytrope region $\mathcal{P}_d$ represent the graduated orders $\Lambda_M \subset K^{d\times d}$. The data above enables us to sample from these orders.

Our next result relates the structure of a polytrope $Q_M$ to that of its graduated order $\Lambda_M$.

**Theorem 5.20.** *Let $M \in \mathcal{P}_d$ be in standard form. The $(d-1)$-dimensional polytrope $Q_M$ is both a min-plus simplex and a max-plus simplex. The min-plus vertices $u$ are the columns of $M$. They represent precisely those modules $L_u$ over the order $\Lambda_M$ that are projective. The max-plus vertices $v$ are the columns of $-M^t$, and they represent the injective $\Lambda_M$-modules $L_v$.*

*Proof.* Thanks to [97, Theorem 7], full-dimensional polytropes are tropical simplices, with vertices given by the columns of the defining matrix $M$. We know from bi-tropical convexity [96, Proposition 5.30] that $Q_M$ is both min-convex and max-convex, so it is a simplex in

both ways. This duality corresponds to swapping $M$ with its negative transpose $-M^t$. Note its appearence in [116, Theorem 5.2.21]. The connection to projective and injective modules appears in parts (v) and (vii) of [133, Remark II.4]. For completeness, we sketch a proof.

Recall that a module is projective if and only if it is a direct summand of a free module. Let $m^{(1)}, \ldots, m^{(d)}$ denote the columns of $M$. The lattice associated to the $j$-th column equals

$$L_{m^{(j)}} \;=\; \big\{\, x \in K^d \,:\, \mathrm{val}(x_i) \geq m_{ij} \;\text{ for } i = 1, \ldots, d \,\big\}.$$

Taking the direct sum of these $d$ lattices gives the following identification of $\mathcal{O}_K$-modules:

$$\Lambda_M \;=\; L_{m^{(1)}} \oplus L_{m^{(2)}} \oplus \cdots \oplus L_{m^{(d)}}. \tag{5.10}$$

We see that $L_{m^{(j)}}$ is a direct summand of the free rank one module $\Lambda_M$, so it is projective.

Conversely, let $P$ be any indecomposable projective $\Lambda_M$-module. Then $P \oplus Q \cong \Lambda_M^r$ for some module $Q$ and some $r \in \mathbb{Z}_{>0}$. The module $\Lambda_M^r$ decomposes into $r \cdot d$ indecomposables, found by aggregating $r$ copies of (5.10). By the Krull-Schmidt Theorem, such decompositions are unique up to isomorphism, and hence $P$ is isomorphic to $L_{m^{(j)}}$ for some $j$.

A $\Lambda_M$-module $P$ is projective if and only if $\mathrm{Hom}_{\mathcal{O}_K}(P, \mathcal{O}_K)$ is an injective $\Lambda_M$-module, but now with the action on the right. The decomposition (5.10) dualizes gracefully. We derive the assertion for injective modules by dualizing all steps in the argument above. □

**Example 5.21.** The columns of the matrix $M$ in Example 5.3 are the negated unit vectors $-e_i$. The columns of $-M^t$ are the unit vectors $e_i$. Our color coding in Figure 5.1 exhibits the two structures of $Q_M$ as a tropical tetrahedron in $\mathbb{R}^4/\mathbb{R}\mathbf{1}$. The four red points are the min-plus vertices, giving the projective $\Lambda_M$-modules. The four blue points are the max-plus vertices.

Given any min-plus idempotent matrix $M \in \mathcal{P}_d$, we define its *truncated polytrope region*

$$\mathcal{P}_d(M) \;=\; \{N \in \mathcal{P}_d \,:\, N \leq M\}. \tag{5.11}$$

This polytope has dimension $d^2 - d$ if $M$ is in the interior of $\mathcal{P}_d$. It parametrizes all subpolytropes of $Q_M$, i.e. all the polytropes $Q_N$ contained in $Q_M$, as the following lemma shows.

**Lemma 5.22.** *Given matrices $M$ in $\mathcal{P}_d$ and $N$ in $\mathbb{R}_0^{d\times d}$ such that $Q_N \subseteq Q_M$, there exists a matrix $C$ in the truncated polytrope region $\mathcal{P}_d(M)$ such that $Q_N = Q_C$.*

*Proof.* For each choice of $i$ and $j$, we define $c_{ij} = \max\{u_i - u_j \,:\, u \in Q_N\}$. The matrix $C = (c_{ij})$ lives in $\mathbb{R}_0^{d\times d}$ and has the property that $Q_N = Q_C$. Moreover, since $Q_N$ is contained in $Q_M$, we have $C \leq M$. The fact that $C \odot C = C$ follows from the definition of the $c_{ij}$'s and (5.4). In particular, $C$ belongs to the truncated polytrope region $\mathcal{P}_d(M)$. □

On the algebraic side, $\mathcal{P}_d(M)$ parametrizes all $\mathcal{O}_K$-orders $\Lambda_N$ that contain the given order $\Lambda_M$. Here $M$ and $N$ are assumed to be integer matrices. In particular, the integer points $u$ in $Q_M$ correspond to maximal orders $\Lambda_{M(u)} = \mathrm{End}_{\mathcal{O}_K}(L_u)$ that contain $\Lambda_M$; cf. Proposition 5.4.

**Example 5.23.** Let $M$ be the $d \times d$ matrix with entries 0 on the diagonal and 1 off the diagonal. Thus $Q_M$ is the pyrope [97, §3]. We consider two cases: the hexagon ($d = 3$) and Example 5.3 ($d = 4$). The truncated polytrope region $\mathcal{P}_d(M)$ classifies subpolytropes of $Q_M$.

$\underline{d = 3}$: The 6-dimensional polytope $\mathcal{P}_3(M)$ has the f-vector $(36, 132, 199, 151, 60, 12)$. Its 36 vertices come in ten symmetry classes. We list the corresponding $3 \times 3$ matrices:

$$\begin{array}{llllllllll} 1,6 & [1,1,1,1,1,1] & 2,6 & [1,\tfrac{1}{2},\tfrac{1}{2},1,1,\tfrac{1}{2}] & 3,8 & [0,-1,0,-1,1,1] & 3,8 & [1,0,-1,-1,0,1] & 3,8 & [1,0,1,1,0,1] \\ 3,6 & [1,1,1,1,0,0] & 3,6 & [0,1,1,1,1,0] & 6,7 & [0,-1,1,0,1,1] & 6,7 & [1,1,1,1,0,1] & 6,6 & [0,0,1,1,1,0] \end{array}$$

These polytropes are shown in red in Figure 5.3. Our classification into $S_3$-orbits is finer than that from symmetries of the hexagon $Q_M$, which leads to only eight orbits. For us, this classification is more natural because it reflects algebraic properties of orders. It distinguishes min-plus vertices from max-plus vertices of $Q_M$. The polytope $\mathcal{P}_3(M)$ has 41 integer points, so there are 41 orders containing $\Lambda_M$. In addition to 34 integer vertices, there are seven interior integer points, namely $[0, 0, 0, 0, 0, 0]$ and six like $[0, 0, 0, 0, 1, 1]$, not seen in Figure 5.3.

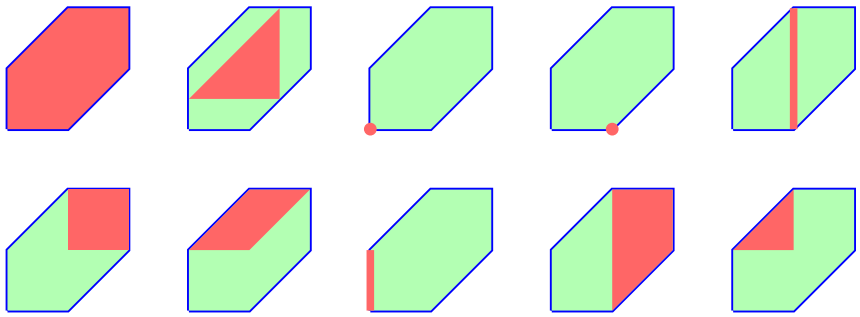


Figure 5.3: The regular hexagon has 36 extreme subpolytropes in ten symmetry classes.

$\underline{d = 4}$: The truncated polytrope region $\mathcal{P}_4(M)$ for (5.1) is 12-dimensional. Its f-vector is

$$(961, 17426, 103780, 304328, 517293, 549723, 377520, 168720, 48417, 8620, 894, 48).$$

The 961 vertices come in 65 orbits under the $S_4$-action. Among the simple vertices we find:

$$\begin{array}{llll} 1,12 & [1,1,1,1,1,1,1,1,1,1,1,1] & 8,12 & [1,1,1,\tfrac{1}{2},1,1,\tfrac{1}{2},1,\tfrac{1}{2},1] \\ 4,27 & [1,1,1,-1,0,0,-1,0,0,-1,0,0] & 4,27 & [-1,-1,-1,1,0,0,1,0,0,1,0,0] \end{array}$$

The list of all vertices, and much more, is available at

https://mathrepo.mis.mpg.de/OrdersPolytropes/index.html.

Such data sets can be useful for comprehensive computational studies of $\mathcal{O}_K$-orders in $K^{d\times d}$.

### 5.2.4 Ideals

To better understand the order $\Lambda_M$ for $M \in \mathcal{P}_d$, we study its (fractional) ideals. By an *ideal* of $\Lambda_M$ we mean an additive subgroup $I$ of $\Lambda_M$ such that $\Lambda_M I \subseteq I$ and $I\Lambda_M \subseteq I$. A *fractional ideal* of $\Lambda_M$ is a (two sided) $\Lambda_M$-submodule $J$ of $K^{d\times d}$ such that $\alpha J \subset \Lambda_M$ for some $\alpha \in K^*$.

**Example 5.24.** Fix $X \in K^{d\times d}$ and consider the two-sided $\Lambda_M$-module $\langle X\rangle = \Lambda_M X \Lambda_M = \{AXB : A, B \in \Lambda_M\}$. This is an ideal when $X \in \Lambda_M$. If $X \notin \Lambda_M$ then $\alpha X \in \Lambda_M$ for some $\alpha \in K^*$. Hence, $\langle X\rangle$ is a fractional ideal. These are the *principal (fractional) ideals* of $\Lambda_M$.

For all that follows, we assume that $M \in \mathcal{P}_d$ is an integer matrix in standard form.

**Proposition 5.25.** *The nonzero fractional ideals of the order $\Lambda_M$ are the sets of the form*

$$I_N \;=\; \left\{ X \in K^{d\times d} \colon\ \mathrm{val}(X) \geq N \right\}, \tag{5.12}$$

*where $N = (n_{ij})$ is any matrix in $\mathbb{Z}^{d\times d}$ with $N \odot M = M \odot N = N$. This is equivalent to*

$$n_{ik} \leq n_{ij} + m_{jk} \quad \text{and} \quad n_{ik} \leq m_{ij} + n_{jk} \quad \text{for} \quad 1 \leq i, j, k \leq d. \tag{5.13}$$

*Proof.* The result is due to Plesken who states it in (viii) from [133, Remark II.4]. The min-plus matrix identity $N \odot M = N$ is equivalent to $n_{ik} \leq n_{ij} + m_{jk}$ because $m_{jj} = 0$. □

**Remark 5.26.** If $N$ has zeros on its diagonal and satisfies (5.4) then $I_N = \Lambda_N$ is an order, as before. However, among all lattices in $K^{d\times d}$, ideals are more general than orders. In particular, we generally have $n_{ii} \neq 0$ for the matrices $N$ in (5.12). A fractional ideal $I_N$ is an ideal in $\Lambda_M$ if and only if $N \geq M$. If this holds then the polytrope $Q_N$ is contained in $Q_M$.

**Example 5.27.** The Jacobson radical of the order $\Lambda_M$ is the ideal $\mathrm{Jac}(\Lambda_M) = I_{M+\mathrm{Id}_d}$. Here $\mathrm{Id}_d$ is the identity matrix. The quotient of $\Lambda_M$ by its Jacobson radical is the product of residue fields $\Lambda_M/\mathrm{Jac}(\Lambda_M) \cong (\mathcal{O}_K/\langle\varpi\rangle)^d$. See (i) in [133, Remark II.4] for more details.

Let $\mathcal{Q}_M$ denote the set of matrices $N$ in $\mathbb{R}^{d\times d}$ that satisfy the inequalities in (5.13). These inequalities are bounds on differences of matrix entries in $N$. We can thus regard $\mathcal{Q}_M$ as a polytrope in $\mathbb{R}^{d\times d}/\mathbb{R}\mathbf{1}$, where $\mathbf{1} = \sum_{i,j=1}^{d} E_{ij}$. The matrices $N$ parameterizing the fractional ideals $I_N$ of $\Lambda_M$ (up to scaling) are the integer points of $\mathcal{Q}_M$. One checks directly that $\mathcal{Q}_M$ is closed under both addition and multiplication of matrices in the min-plus algebra. Its product $\odot$ represents the multiplication of fractional ideals as the following proposition shows.

**Proposition 5.28.** *If $M \in \mathcal{P}_d$ is in standard form and $N, N' \in \mathcal{Q}_M$ then $I_N I_{N'} = I_{N \odot N'}$.*

*Proof.* Let $X \in I_N, Y \in I_{N'}$. The inequalities $\mathrm{val}(X) \geq N$, $\mathrm{val}(Y) \geq N'$ imply $\mathrm{val}(XY) \geq \mathrm{val}(X) \odot \mathrm{val}(Y) \geq N \odot N'$ and so $XY \in I_{N\odot N'}$. This gives the inclusion $I_N I_{N'} \subseteq I_{N\odot N'}$. Let $u_{ij} = \min_{1\leq k\leq d}(n_{ik} + n'_{kj})$ be the $(i,j)$ entry of $N \odot N'$. For the inclusion $I_{N\odot N'} \subseteq I_N I_{N'}$, it suffices to show that $\varpi^{u_{ij}} E_{ij}$ is in $I_N I_{N'}$ for all $i, j$. Fix $i, j$ and let $k$ satisfy $u_{ij} = n_{ik} + n'_{kj}$. The matrices $\varpi^{n_{ik}} E_{ik}$ and $\varpi^{n'_{kj}} E_{kj}$ are in $I_N$ and $I_{N'}$. Their product $\varpi^{u_{ij}} E_{ij}$ is in $I_N I_{N'}$. □

We call $\mathcal{Q}_M$ the *ideal class polytrope* of $M$. The min-plus semigroup $(\mathcal{Q}_M, \underline{\odot})$ plays the role of the ideal class group in number theory. Its neutral element is the given matrix $M$.

**Example 5.29.** Fix $M = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \in \mathcal{P}_2$. The polytrope $\mathcal{Q}_M$ is the octahedron with vertices

$$\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \in \mathbb{Z}^{2\times 2}/\mathbb{Z}\mathbf{1}.$$

This octahedron contains 19 integer points $N$. These are in bijection with the equivalence classes of fractional ideals $I_N$ in the order $\Lambda_M$. The midpoint of $\mathcal{Q}_M$ corresponds to the Jacobson radical $I_{M+\mathrm{Id}_2}$. The remaining 12 integer points are the midpoints of the edges.

One may ask whether the ideal class semigroup $(\mathcal{Q}_M, \underline{\odot})$ is actually a group. To address this question, we define the *pseudo-inverse* of a fractional ideal $I$ in the order $\Lambda_M$ as follows:

$$(\Lambda_M : I) \;=\; \{\, X \in K^{d\times d} : XI \subseteq \Lambda_M \;\text{ and }\; IX \subseteq \Lambda_M \}.$$

**Lemma 5.30.** *The pseudo-inverse of a fractional ideal in $\Lambda_M$ is a fractional ideal in $\Lambda_M$.*

*Proof.* Let $A \in \Lambda_M$ and $X \in (\Lambda_M : I)$, so that $XI, IX \subseteq \Lambda_M$. Since $I$ is a fractional ideal, we have $AI \subseteq I$ and $IA \subseteq I$. From these inclusions we deduce that $XAI, IXA, AXI, IAX$ are all subsets of $\Lambda_M$. This implies $XA, AX \in (\Lambda_M : I)$. Hence $(\Lambda_M : I)$ is a fractional ideal. □

**Proposition 5.31.** *Let $M \in \mathcal{P}_d$ in standard form and $N \in \mathcal{Q}_M$. Then $(\Lambda_M : I_N) = I_{N'}$ where*

$$n'_{ij} \;=\; \max_{1\le \ell \le d} \big( \max(m_{\ell j} - n_{\ell i}, m_{i\ell} - n_{j\ell}) \big) \qquad \textit{for} \quad 1 \le i, j \le d. \tag{5.14}$$

*Proof.* By Proposition 5.25 and Lemma 5.30, there exists $N' \in \mathcal{Q}_M$ with $I_{N'} = (\Lambda_M : I_N)$. Then $I_{N'} I_N \subseteq \Lambda_M$ and $I_N I_{N'} \subseteq \Lambda_M$, and $I_{N'}$ is the largest fractional ideal with this property. These two conditions are equivalent to $\varpi^{n'_{ij}} E_{ij} I_N \subseteq \Lambda_M$ and $\varpi^{n'_{ij}} I_N E_{ij} \subseteq \Lambda_M$ for all $i, j$. The first condition holds if and only if $n'_{ij} + n_{j\ell} \ge m_{i\ell}$ for all $\ell$. The second condition holds if and only if $n_{\ell i} + n'_{ij} \ge m_{\ell j}$ for all $\ell$. The smallest solution $N' = (n'_{ij})$ is given by (5.14). □

Passing from ideals to their matrices, we also call $N'$ the *pseudo-inverse* of $N$ in $\mathcal{Q}_M$.

**Example 5.32.** Let $d = 2$ and $M$ as in Example 5.29. The 19 ideal classes $N$ in $\mathcal{Q}_M$ have only three distinct pseudo-inverses: $N' \in \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$. For most ideal classes $N$, we have $N \underline{\odot} N' \neq M$ and $N' \underline{\odot} N \neq M$. This means that most $N$ do not have an inverse in $(\mathcal{Q}_M, \underline{\odot})$. In particular, the ideal class polytrope $\mathcal{Q}_M$ is a semigroup but not a group.

The semigroup $\mathcal{Q}_M$ has the neutral element $M$ and each ideal class $N \in \mathcal{Q}_M$ has a pseudo-inverse $N'$ given by the formula (5.14). With this data, we define the *ideal class group*

$$\mathcal{G}_M \;=\; \{\, N \in \mathcal{Q}_M \,:\, N \underline{\odot} N' \,=\, N' \underline{\odot} N \,=\, M \}\,.$$

This is the maximal subgroup of the semigroup $\mathcal{Q}_M$. It would be interesting to understand how $M$ determines the structure of $\mathcal{G}_M$. Note that $\mathcal{G}_M = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ in Example 5.32.

**Example 5.33.** Here are three examples of ideal class groups of graduated orders:

$$J_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad J_3 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad J_4 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$
$$\mathcal{G}_{J_2} \cong \mathbb{Z}/2\mathbb{Z} \qquad \mathcal{G}_{J_3} \cong \mathbb{Z}/6\mathbb{Z} \qquad \mathcal{G}_{J_4} \cong S_4$$

How does this list continue as we pass from Example 5.3 to pyropes [97, §3] in higher dimensions?

We end this section with a conjecture about the geometry of $\mathcal{G}_M$ inside $\mathcal{Q}_M$.

**Conjecture 5.34.** For any integer matrix $M$ in the polytrope region $\mathcal{P}_d$, the elements in the ideal class polytrope $\mathcal{G}_M$ are among the classical vertices of the ideal class polytrope $\mathcal{Q}_M$.

### 5.2.5 Towards the Bruhat-Tits Building

Affine buildings [1, 169] provide a natural setting for orders and min-max convexity. The objects we discussed in this chapter so far are associated to one apartment in this building, namely the apartment corresponding to the diagonal lattices. The aim of this section is to present this perspective and to lay the foundation for the following section where we discuss *bolytropes* and *bolytrope orders*.

We start by recalling reviewing some notions from Section 1.2.

**Definition 5.35.** The *affine building* $\mathcal{B}_d(K)$ is an infinite simplicial complex. Its vertices are the equivalence classes $[L]$ of lattices in $K^d$. A configuration $\{[L_1], \ldots, [L_s]\}$ is a simplex in $\mathcal{B}_d(K)$ if and only if, up to some permutation, there exist representatives $\tilde{L}_i \in [L_i]$ satisfying $\tilde{L}_1 \supset \tilde{L}_2 \supset \cdots \supset \tilde{L}_s \supset \varpi \tilde{L}_1$. The maximal simplices $\{[L_1], \ldots, [L_d]\}$ are called *chambers*. The *standard chamber* $C_0$ is given by the diagonal lattices $L_i = L_{(\mathbf{1}_{i-1}, \mathbf{0}_{d-i+1})} = L_{(1,\ldots,1,0,\ldots,0)}$.

Given a basis $\{b_1, \ldots, b_d\}$ of $K^d$, the *apartment* defined by this basis is the set of classes $[L]$ of all lattices $L = \bigoplus_{i=1}^d \varpi^{u_i} \mathcal{O}_K b_i$ where $u_1, \ldots, u_d$ range over $\mathbb{Z}$. Hence the apartment is

$$\{ [\varpi^{u_1} \mathcal{O}_K b_1 \oplus \cdots \oplus \varpi^{u_d} \mathcal{O}_K b_d] : u_1, \ldots, u_d \in \mathbb{Z} \} \;=\; \{ [g L_u] : u \in \mathbb{Z}^d \},$$

where $g \in \mathrm{GL}(d, K)$ is the matrix with columns $b_1, \ldots, b_d$. The *standard apartment* is the one associated with the standard basis $(e_1, \ldots, e_d)$ of $K^d$. The vertices of the standard apartment are the diagonal lattice classes $[L_u]$ for $u \in \mathbb{Z}^d$. We identify this set of vertices with $\mathbb{Z}^n/\mathbb{Z}\mathbf{1}$.

The general linear group $\mathrm{GL}_d(K)$ acts on the building $\mathcal{B}_d(K)$. This action preserves the simplicial complex structure. In fact, the action is transitive on lattice classes, on apartments and also on the chambers. The stabilizer of the standard lattice $L_0$ is the subgroup

$$\mathrm{GL}(d, \mathcal{O}_K) \;=\; \{ g \in \mathcal{O}_K^{d \times d} \,:\, \mathrm{val}(\det(g)) = 0 \} \;\subset\; \mathrm{GL}(d, K).$$

Starting from the standard chamber $C_0$, there exist reflections $s_0, s_1, \ldots, s_{d-1}$ in $\mathrm{GL}(d, K)$ that map $C_0$ to the $d$ adjacent chambers in the standard apartment. For $i \geq 1$, define $s_i$ by

$$s_i(e_i) = e_{i+1}, \quad s_i(e_{i+1}) = e_i \text{ and } s_i(e_j) = e_j \text{ when } j \neq i, i+1.$$

The map $s_0$ is defined by $s_0(e_i) = e_i$ for $i = 2, \ldots, d-1$ and $s_0(e_d) = pe_1$, $s_0(e_1) = \varpi^{-1}e_d$. The reflections $s_0, \ldots, s_{d-1}$ are Coxeter generators for the *affine Weyl group* $W = \langle s_0, \ldots, s_{d-1} \rangle$. The group $W$ acts regularly on the chambers $C$ in the standard apartment [22, § 1.5, Thm. 2]: for every $C$ there is a unique $w \in W$ such that $C = wC_0$. The elements of $W$ are the matrices $h_\sigma g_u$ where $h_\sigma = (1_{i=\sigma(j)})_{i,j}$ for $\sigma \in S_d$, and $u \in \mathbb{Z}^d$ with $u_1 + \cdots + u_d = 0$. Thus $W$ is the semi-direct product of $S_d$ and the group of diagonal matrices $g_u$ whose exponents sum to 0.

Our primary object of interest is the Plesken-Zassenhaus order $\mathrm{PZ}(\Gamma)$ of a finite configuration $\Gamma$ in the affine building $\mathcal{B}_d(K)$. This is the intersection (5.8) of endomorphism rings. In this section we study the case when $\Gamma$ lies in one apartment. In Theorem 5.14 we showed that $\mathrm{PZ}(\Gamma) = \Lambda_M$ where $M$ is the matrix in $\mathcal{P}_d$ that encodes the min-max convex hull of $\Gamma$. This was used in Sections 5.2.3 and 5.2.4 to elucidate combinatorial and algebraic structures in $\mathrm{PZ}(\Gamma)$.

The min-max convex hull of $\Gamma$ will play the role of the polytrope, with its distinguished vertices representing injective and projective modules, as in Theorem 5.20. Computing that convex hull requires tools as in [169] but simultaneously in min-plus algebra and max-plus algebra.

We conclude this section with configurations given by two chambers $C, C'$ in $\mathcal{B}_d(K)$. We are interested in the their order $\mathrm{PZ}(C \cup C')$. A fundamental fact about buildings states that any two chambers $C, C'$ lie in a common apartment, cf. [22, 1]. Also, since the affine Weyl group $W$ acts regularly on the chambers of the standard apartment, we can then reduce to the case where the two chambers in question are $C_0$ and $wC_0$ for some $w = h_\sigma g_u \in W$.

**Example 5.36.** The standard chamber $C_0$ is encoded by $M_0 = \sum_{1 \leq i < j \leq d} E_{ij}$. The polytrope $Q_{M_0}$ is a simplex. The order $\mathrm{PZ}(C_0) = \Lambda_{M_0}$ consists of all $X \in \mathcal{O}_K^{d \times d}$ with $x_{ij} \in \langle \varpi \rangle$ for $i < j$.

Let $D_u = \mathrm{val}(g_u)$ denote the tropical diagonal matrix with $u_1, \ldots, u_d$ on the diagonal and $+\infty$ elsewhere. We also write $P_\sigma := \mathrm{val}(h_\sigma)$ for the tropical permutation matrix given by $\sigma$.

**Proposition 5.37.** *We have* $\mathrm{PZ}(C_0 \cup h_\sigma g_u C_0) = \Lambda_{M^{\sigma,u}}$ *where the matrix* $M^{\sigma,u}$ *is given by*

$$M^{\sigma,u} \;=\; M_0 \;\overline{\oplus}\; (P_\sigma \odot D_u \odot M_0 \odot D_{-u} \odot P_{\sigma^{-1}})\,.$$

*Proof.* We have $\mathrm{PZ}(C_0 \cup h_\sigma g_u C_0) = \mathrm{PZ}(C_0) \cap \mathrm{PZ}(h_\sigma g_u C_0)$. Recall that $\mathrm{PZ}(C_0) = \Lambda_{M_0}$ from Example 5.36. Suppose that $M \in \mathbb{Z}_0^{d \times d}$ satisfies $\mathrm{PZ}(h_\sigma g_u C_0) = \Lambda_M$. By Theorem 5.14, the order $\Lambda_{M_0 \,\overline{\oplus}\, M}$ is equal to $\mathrm{PZ}(C_0 \cup h_\sigma g_u C_0)$. To determine $M$, notice that $\mathrm{PZ}(wC_0) = h_\sigma g_u \,\mathrm{PZ}(C_0) g_{-u} h_{\sigma^{-1}}$. This implies the stated formula $M = P_\sigma \odot D_u \odot M_0 \odot D_{-u} \odot P_{\sigma^{-1}}$. □

We may ask for invariants of the orders $\mathrm{PZ}(C_0 \cup wC_0)$ in terms of $w \in W$. Clearly, not all polytropes in an apartment arise as the min-max convex hull of two chambers. Which graduated orders are of the form $\mathrm{PZ}(C_0 \cup wC_0)$? Which other elements $w'$ in the affine Weyl group $W$ give rise to the same Plesken-Zassenhaus order $\mathrm{PZ}(C_0 \cup wC_0)$ up to isomorphism?

## 5.3 Bolytrope orders

In this section, we extend our study to a bigger class of orders we call *bolytrope orders*. Roughly speaking, we shall see that these orders arise as the Plesken Zassenhaus rings of the Minkowski sum of a polytrope and a ball in the building $\mathcal{B}_d(K)$. Hence the name *bolytrope* which is derived from *ball* and *polytrope*. We will start by defining a distance on $\mathcal{B}_d^0(K)$ (the set of 0-simplices in $\mathcal{B}_d(K)$) and use it to define balls and bolytropes in the building $\mathcal{B}_d(K)$. Balls are a special type of bolytropes and bolytropes can be thought of as balls "around polytropes".

As before, we denote by $\mathbf{1}$ the vector $(1,\dots,1) \in \mathbb{Z}^d$ and by $J_d$ the matrix, in $\mathbb{Z}^{d\times d}$, with zeros on the diagonal and ones elsewhere.

### 5.3.1 The distance

The content of this section heavily depends on the following notion of distance on $\mathcal{B}_d^0(K)$.

**Definition 5.38.** Let $[L_1],[L_2] \in \mathcal{B}_d^0(K)$ be two homothety classes of lattices. Then

$$\operatorname{dist}([L_1],[L_2]) := \min\{s\colon \text{ there are } L_1' \in [L_1], L_2' \in [L_2] \text{ with } \varpi^s L_1' \subseteq L_2' \subseteq L_1'\}.$$

For a subset $\mathcal{L} \subseteq \mathcal{B}_d^0(K)$, we put $\operatorname{dist}([L],\mathcal{L}) := \min\{\operatorname{dist}([L],[L'])\colon [L'] \in \mathcal{L}\}$ and

$$\operatorname{diam}(\mathcal{L}) := \sup_{[L],[L']\in\mathcal{L}} \operatorname{dist}([L],[L'])$$

The set $\mathcal{L}$ is called *bounded*, if its *diameter* $\operatorname{diam}(\mathcal{L})$ is finite.

The following result justifies the name distance for dist.

**Lemma 5.39.** *The map* $\operatorname{dist} : \mathcal{B}_d^0(K) \times \mathcal{B}_d^0(K) \to \mathbb{Z}$ *is a distance on* $\mathcal{B}_d^0(K)$.

*Proof.* We check that the defining properties of a distance hold. For this, let $[L_1],[L_2] \in \mathcal{B}_d^0(K)$ with $\operatorname{dist}([L_1],[L_2]) = s$ and let $L_1', L_2'$ be as in Definition 5.38. Then

(1) $\operatorname{dist}([L_1],[L_2]) = 0$ if and only $L_1' \subseteq L_2' \subseteq L_1'$, equivalently $[L_1] = [L_2]$.

(2) If $\varpi^s L_1' \subseteq L_2' \subseteq L_1$ then $\varpi^s L_2' \subseteq \varpi^s L_1' \subseteq L_2'$, so $\operatorname{dist}([L_1],[L_2]) = \operatorname{dist}([L_2],[L_1])$.

(3) Let $[L_3] \in \mathcal{B}_d^0(K)$ and set $s' = \operatorname{dist}([L_2],[L_3])$. Let, moreover $L_3' \in [L_3]$ and $L_2'' \in [L_2]$ be such that $\varpi^{s'} L_2'' \subseteq L_3' \subseteq L_2''$. Write $L_2'' = \varpi^t L_2'$. Then

$$\varpi^t L_1' \supseteq \varpi^t L_2' \supseteq L_3' \supseteq \varpi^{s'+t} L_2' \supseteq \varpi^{s+s'+t} L_1' = \varpi^{s+s'}(\varpi^t L_1'),$$

yielding that $\operatorname{dist}([L_1],[L_2]) + \operatorname{dist}([L_2],[L_3]) \geq \operatorname{dist}([L_1],[L_3])$.

The choices of $[L_1],[L_2],[L_3]$ being arbitrary, the proof is complete. □

Thanks to the elementary divisor theorem for modules over PIDs, we know that any two lattices in $K^d$ have compatible bases, i.e. for any two lattice classes $[L_1]$ and $[L_2]$, there is always an apartment containing both. So, to compute their distance, we may choose a frame basis $(e_1, \dots, e_d)$ of $K^d$, so that $L_1 = L_{(0,\dots,0)}$ and $L_2 = L_{(u_1,\dots,u_d)}$ with $u_1 \geq \dots \geq u_d$. With this choice, we obtain that $\operatorname{dist}([L_1], [L_2]) = u_1 - u_d$.

**Remark 5.40.** The distance between lattice classes $[L_u]$ and $[L_v]$ in the same apartment $\mathcal{A}(E)$ is given by

$$\operatorname{dist}([L_u], [L_v]) = \max_{1 \leq i \leq d}(v_i - u_i) - \min_{1 \leq j \leq d}(v_j - u_j).$$

In particular, any bounded subset of an apartment is finite. For a connection to tropical geometry, see for instance [96, Section 5.3].

Note that the distance from Definition 5.38 coincides with the 1-skeleton distance on $\mathcal{B}_d^0(K)$, as the following result shows. For $L$ and $L'$ lattices with $\varpi L \subset L' \subset L$, write $([L], [L'])$ for the 1-simplex with ends $[L]$ and $[L']$.

**Lemma 5.41.** *Let $[L_1], [L_2] \in \mathcal{B}_0^d(K)$ be distinct and set $s = \operatorname{dist}([L_1], [L_2])$. Then $s > 0$,*

(1) *there exist $[L_1] = [X_0], [X_1] \dots, [X_{s-1}], [X_s] = [L_2] \in \mathcal{B}_d^0(K)$ such that $([X_{i-1}], [X_i])$ are 1-simplices for all $1 \leq i \leq s$, and*

(2) *there is no shorter sequence connecting $[L_1]$ and $[L_2]$ in the 1-skeleton of $\mathcal{B}_d(K)$.*

*Proof.* The number $s$ is positive as a consequence of Lemma 5.39. Without loss of generality, assume that $\varpi^s L_1 \subseteq L_2 \subseteq L_1$ and put $X_1 := \varpi L_1 + L_2$. Then $\varpi L_1 \subseteq X_1 \subseteq L_1$ and so $([L_1], [X_1])$ is a 1-simplex in $\mathcal{B}_d(K)$. For $i = 2, \dots, s$, put $X_i := \varpi X_{i-1} + L_2 = \varpi^i L_1 + L_2$. Then $X_s = L_2$ and all $([X_{i-1}], [X_i])$ are 1-simplices in the building. We have proven (1), while (2) follows from the triangle inequality and the fact that two lattice classes in a 1-simplex have distance at most 1. □

### 5.3.2 Balls and bolytropes

**Definition 5.42.** Let $\mathcal{L}$ be a bounded subset of $\mathcal{B}_d^0(K)$. Then the *closed ball of radius $r$ and center $\mathcal{L}$* is

$$\mathrm{B}_r(\mathcal{L}) := \{[L] \in \mathcal{B}_d^0(K) \colon \operatorname{dist}([L], \mathcal{L}) \leq r\}.$$

If $\mathcal{L} = \{[L]\}$ consists of one element only, then

$$\mathrm{B}_r([L]) := \mathrm{B}_r(\mathcal{L})$$

is the *ball with center $[L]$ and radius $r$*. If $\mathcal{L} = Q(\Lambda_M)$, then

$$\mathrm{B}_r(M) := \mathrm{B}_r(Q(\Lambda_M))$$

is called the *bolytrope with center $Q(\Lambda_M)$* and radius $r$.

In particular, the ball $\mathrm{B}_r([L])$ consists of all lattice classes $[L']$ that are represented by some lattice $L'$ such that $\varpi^r L \subseteq L' \subseteq L$. We close the section by computing the intersection of a bolytrope with an apartment. Recall that $J_d \in \mathcal{P}_d(\mathbb{Z})$ is the matrix with all 1s outside of the main diagonal.

**Lemma 5.43.** *Let $\mathcal{A}$ be an apartment containing $Q(\Lambda_M)$. Then*

$$\mathrm{B}_r(M) \cap \mathcal{A} = Q(\Lambda_{M+rJ_d}).$$

*Proof.* Let $(e_1, \ldots, e_d)$ be a frame basis defining $\mathcal{A}$ and put $Q = Q(\Lambda_{M+rJ_d})$. We will use Remark 5.11 with respect to this basis. Since $\varpi^r \Lambda_M \subseteq \Lambda_{M+rJ_d} \subseteq \Lambda_M$, we have the inclusion $Q \subseteq \mathrm{B}_r(M)$. Now we show the other inclusion. Let $[L_u]$ in $\mathcal{A}$ be of distance at most $r$ from some lattice $[L_v] \in Q(\Lambda_M)$. Suppose that $[L_u] \notin Q(\Lambda_{M+rJ_d})$. This means that there exist $1 \leq i \neq j \leq d$ such that $u_i - u_j > m_{ij} + r$. However, since $[L_v] \in Q(\Lambda_M)$, we have $v_i - v_j \leq m_{ij}$ and hence $u_i - u_j > v_i - v_j + r$. In other words

$$(u_i - v_i) - (u_j - v_j) > r, \text{ so } \operatorname{dist}([L_u], [L_v]) > r.$$

This is a contradiction and so the proof is complete. □

#### 5.3.2.1 Plesken-Zassenhaus closed sets.

We have seen that closed orders are determined by the collection of their stable lattices; such sets are thus of fundamental importance for the study of closed orders.

**Definition 5.44.** A subset $\mathcal{L}$ of $\mathcal{B}_d^0(K)$ is called PZ-*closed* if $\mathcal{L} = Q(\Lambda)$ for some order $\Lambda$.

For the study of PZ-closed subsets it clearly suffices to consider closed orders $\Lambda$. Note that the bijection $\Lambda \leftrightarrow Q(\Lambda)$ is a Galois correspondence between

$$\{ \text{ closed orders in } K^{d\times d} \} \longleftrightarrow \{ \text{ PZ-closed subsets of } \mathcal{B}_d^0(K) \}.$$

**Remark 5.45.** Let $M := \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right) \in \mathcal{P}_d(\mathbb{Z})$. Then $\Lambda_M$ is a graduated order with

$$Q(\Lambda_M) = \{[L_{(0,1)}], [L_{(0,0)}]\}.$$

Let $\Lambda := \{X \in \Lambda_M \colon X_{11} \equiv X_{22} \bmod \varpi\}$. Then $\Lambda$ is an order in $K^{2\times 2}$ satisfying $Q(\Lambda) = Q(\Lambda_M)$. It follows that $\Lambda$ is not a closed order.

**Remark 5.46.** The following is a summary of Proposition 5.8, Corollary 5.13 and Theorem 5.20. If $M \in \mathcal{P}_d(\mathbb{Z})$, then the graduated order $\Lambda_M$ is closed and

$$Q(\Lambda_M) = \{[L_u] \colon u \in \mathbb{Z}^d,\ u + \mathbb{R}\mathbf{1} \in Q_M\}$$

is a finite set which we can identify with the integral points of the polytrope $Q_M$. Moreover, the projective $\Lambda_M$-lattices are given by the columns of $M$ in the following way: if

$M^{(1)}, \dots, M^{(d)}$ denote the columns of $M$, then, for each projective $\Lambda_M$-lattice $L$, there exists $i \in \{1, \dots, d\}$ such that $L$ is homothetic to

$$P_i := \Lambda_M \epsilon_i = L_{M^{(i)}}.$$

The polytrope $Q_M$ is the min-convex hull of the set $\{M^{(1)} + \mathbb{R}\mathbf{1}, \dots, M^{(d)} + \mathbb{R}\mathbf{1}\}$ and has dimension $\dim(Q_M) = |\{[P_1], \dots, [P_d]\}| - 1$. The order $\Lambda_M = \mathrm{PZ}([P_1], \dots, [P_d])$ is the Plesken-Zassenhaus order of its projective (i.e. projective as $\Lambda_M$-modules) lattices in $K^d$.

As noted in Remark 5.46, the PZ-closed subsets of one apartment $\mathcal{A}$ are exactly the finite and convex subsets of $\mathcal{B}_d^0(K)$, i.e. the polytropes. In general, being bounded and convex is a necessary but not sufficient condition for a subset of $\mathcal{B}_d^0(K)$ to be closed.

**Proposition 5.47.** *Let $\Lambda$ be an order in $K^{d\times d}$. Then $Q(\Lambda)$ is a non-empty bounded convex subset of $\mathcal{B}_d^0(K)$.*

*Proof.* As any order is contained in a maximal order, there is some maximal order $\Gamma$, with $\Lambda \subseteq \Gamma$. Both lattices $\Lambda$ and $\Gamma$ have full rank in $K^{d\times d}$, so there is $r \in \mathbb{Z}_{\geq 0}$ such that $\varpi^r\Gamma \subseteq \Lambda \subseteq \Gamma$. If $[L]$ is the unique class of $\Gamma$-lattices, then $[L] \in Q(\Lambda)$ and hence $Q(\Lambda)$ is not empty. Moreover, all lattice classes in $Q(\Lambda)$ have a representative between $L$ and $(\varpi^r\Gamma)L = \varpi^r L$, so $Q(\Lambda)$ is contained in the ball of radius $r$ around $[L]$. In particular, $Q(\Lambda)$ is bounded. To see convexity, let $[L'], [L''] \in Q(\Lambda)$. Then there is an apartment containing both lattice classes, so $\Gamma' := \mathrm{End}_{\mathcal{O}_K}(L') \cap \mathrm{End}_{\mathcal{O}_K}(L'')$ is a graduated order containing $\Lambda$. But then the convex set $Q(\Gamma') \subseteq Q(\Lambda)$ contains both lattice classes $[L']$ and $[L'']$, and, $[L']$ and $[L'']$ being arbitrary, $Q(\Lambda)$ is convex. □

**Remark 5.48.** Let $\Lambda$ be an order in $K^{d\times d}$ and let $\mathcal{A}$ be an apartment in $\mathcal{B}_d(K)$ such that $Q(\Lambda) \cap \mathcal{A} \neq \emptyset$. Then

$$Q(\Lambda) \cap \mathcal{A} = Q(\Gamma)$$

for a unique graduated overorder $\Gamma$ of $\Lambda$. Indeed, if $\mathcal{A} = \mathcal{A}(E)$ and $\mathcal{E} = \{\epsilon_1, \dots, \epsilon_d\}$ is the set of projections on the frame $E$, then there are only finitely many maximal overorders of $\Lambda$ that contain $\mathcal{E}$. Their intersection is the desired graduated order $\Gamma$.

**Definition 5.49.** Let $\mathcal{L}$ be a bounded subset of $\mathcal{B}_d^0(K)$. The *Plesken-Zassenhaus order* associated to $\mathcal{L}$ is

$$\mathrm{PZ}(\mathcal{L}) := \bigcap_{[L]\in\mathcal{L}} \mathrm{End}_{\mathcal{O}_K}(L).$$

**Proposition 5.50.** *The Plesken-Zassenhaus order $\mathrm{PZ}(\mathcal{L})$ of a bounded subset $\mathcal{L}$ of $\mathcal{B}_d^0(K)$ is an $\mathcal{O}_K$-order in $K^{d\times d}$*

*Proof.* Put $\Lambda = \mathrm{PZ}(\mathcal{L})$. Then $\Lambda$ is an $\mathcal{O}_K$-module that is closed under multiplication and contains $\mathrm{Id}_d$. It remains to show that $\Lambda$ is of full rank in $K^{d\times d}$. As $\mathcal{L}$ is bounded, there are $[L] \in \mathcal{L}$ and $r \in \mathbb{Z}_{\geq 0}$ such that $\mathcal{L} \subseteq \mathbb{B}_r([L])$. For the maximal order $\Gamma = \mathrm{End}_{\mathcal{O}_K}(L)$ we hence have that $\varpi^r\Gamma L' \subseteq L'$ for all $[L'] \in \mathcal{L}$. So $\varpi^r\Gamma \subseteq \Lambda \subseteq \Gamma$ and, as $\varpi^r\Gamma$ contains a $K$-basis of $K^{d\times d}$, the same is true for $\Lambda$. □

The next proposition shows that closed orders are always an intersection of finitely many maximal orders.

**Proposition 5.51.** *Let $\mathcal{L} \subseteq \mathcal{B}_d^0(K)$ be bounded and let $\Lambda = \mathrm{PZ}(\mathcal{L})$ denote its Plesken-Zassenhaus order. Then there exists a finite subset $\{[L_1], \dots, [L_n]\}$ of $\mathcal{L}$ such that $\Lambda = \mathrm{PZ}([L_1], \dots, [L_n])$.*

*Proof.* Choose $[L_1] \in \mathcal{L}$ arbitrarily and put $\Gamma = \mathrm{End}_{\mathcal{O}_K}(L_1)$. As $\mathcal{L}$ is bounded, there is $r \in \mathbb{Z}_{\geq 0}$ such that $\mathcal{L} \subseteq \mathrm{B}_r([L_1])$ and so

$$\varpi^r \Gamma \subseteq \Lambda \subseteq \Gamma.$$

In particular, the $\mathcal{O}_K$-module $\Gamma/\Lambda$ has finite composition length (at most the composition length $d^2 r$ of $\Gamma/\varpi^r\Gamma$). We proceed by induction on this composition length. If $\Gamma = \Lambda$ then we are done, otherwise there is some $[L_2] \in \mathcal{L}$ such that $[L_2] \notin Q(\Gamma)$. Replace $\Gamma$ by $\Gamma \cap \mathrm{End}_{\mathcal{O}_K}(L_2) = \mathrm{PZ}([L_1], [L_2])$ to decrease the composition length of $\Gamma/\Lambda$. After finitely many steps this process constructs the finite set $\{[L_1], \dots, [L_n]\}$ with $\Lambda = \mathrm{PZ}([L_1], \dots, [L_n])$. □

For a closed order $\Lambda$, the minimal cardinality of a set $\mathcal{L}$ such that $\Lambda = \mathrm{PZ}(\mathcal{L})$ is hence an interesting invariant.

**Definition 5.52.** Let $\Lambda$ be a closed order. Then the *degree* of $\Lambda$ is

$$\deg(\Lambda) := \min\{|\mathcal{L}| - 1 \colon \mathcal{L} \subseteq \mathcal{B}_d^0(K) \text{ with } \Lambda = \mathrm{PZ}(\mathcal{L})\}.$$

Thanks to Proposition 5.51, any closed order is a finite intersection of maximal orders, so the degree of a closed order is always finite. The closed orders of degree 0 are exactly the maximal orders and the ones of degree 1 are certain graduated orders. In general, the degree of a graduated order $\Lambda_M$ is equal to $\dim(Q_M)$, cf. Remark 5.46. In the coming sections, we will see that, for ball orders and bolytrope orders, the degree is always bounded from above by $d$, cf. Theorems 5.65 and 5.75, though such a bound need not always be sharp, cf. Remark 5.66.

### 5.3.3 The radical idealizer process

Let $\Lambda$ be an order in $K^{d\times d}$. In this section, we describe the radical idealizer chain of $\Lambda$, a construction that will be at the foundation of the proofs of our main results.

**Definition 5.53.** Let $\Lambda$ and $L$ be an order and a lattice in $K^{d\times d}$, respectively.

- The *Jacobson radical* $\mathrm{Jac}(\Lambda)$ of $\Lambda$ is the intersection of all maximal left ideals of $\Lambda$.
- The *idealizer* of $L$ is $\mathrm{Id}(L) := \{X \in K^{d\times d} \colon XL \subseteq L \text{ and } LX \subseteq L\}$.

**Remark 5.54.** If $\Lambda$ is an order in $K^{d\times d}$, then $\mathrm{Jac}(\Lambda)$ is a two-sided ideal of $\Lambda$ that contains $\varpi\Lambda$. The quotient $\Lambda/\mathrm{Jac}(\Lambda)$ is a semisimple $\mathcal{O}_K/\mathfrak{m}_K$-algebra and, for some $n$, one has $\mathrm{Jac}(\Lambda)^n \subseteq \varpi\Lambda$. Moreover, $\mathrm{Jac}(\Lambda)$ is the unique pro-nilpotent ideal with semisimple quotient ring. For this and more, see for instance [138, Chapter 1, Section 6].

**Definition 5.55.** Let $\Lambda$ be an order in $K^{d\times d}$. The *radical idealizer chain* $(\Omega_i)_{i\geq 0}$ of $\Lambda$ is recursively defined by

$$\Omega_0 := \Lambda \text{ and } \Omega_{i+1} = \mathrm{Id}(\mathrm{Jac}(\Omega_i)).$$

**Remark 5.56.** The radical idealizer chain of an order $\Lambda$ is an ascending finite chain $\Omega_0 \subset \Omega_1 \subset \ldots \subset \Omega_s = \Omega_{s+1} = \ldots$; cf. [129, Remark 3.8]. Moreover, as $\varpi\Lambda \subseteq \mathrm{Jac}(\Lambda)$, we have

$$\Lambda \subseteq \Omega_1 = \mathrm{Id}(\mathrm{Jac}(\Lambda)) \subset \frac{1}{\varpi}\Lambda.$$

This yields an algorithm to compute the radical idealizer chain for orders based on solving linear equations in the residue field; cf. [129]. The sets of invariant lattices $\mathcal{L}_i := Q(\Omega_i)$ form a descending chain

$$\mathcal{L}_0 \supset \mathcal{L}_1 \supset \ldots \supset \mathcal{L}_s,$$

where the last element $\mathcal{L}_s = Q(\Omega_s)$ is known to be a simplex in the building $\mathcal{B}_d(K)$; cf. [138, Theorem (39.14)]. The length $s \geq 0$ of the radical idealizer chain is called the *radical idealizer length* of the order $\Lambda$.

**Lemma 5.57.** *Let $\Lambda$ be an order in $K^{d\times d}$ and put $\Omega_1 := \mathrm{Id}(\mathrm{Jac}(\Lambda))$. Then*

$$Q(\Omega_1) \subseteq Q(\Lambda) \subseteq \mathbb{B}_1(Q(\Omega_1)).$$

*In particular, all lattices in $Q(\Lambda)$ have distance at most one from $Q(\Omega_1)$.*

*Proof.* As $\Omega_1 \supseteq \Lambda$, we know that $Q(\Omega_1) \subseteq Q(\Lambda)$ and thus we get $Q(\Omega_1) = \{[\Omega_1 L] \colon [L] \in Q(\Lambda)\}$. Moreover, by Remark 5.56, we have $\Lambda \subseteq \Omega_1 \subseteq \frac{1}{\varpi}\Lambda$, so $L \subseteq \Omega_1 L \subseteq \frac{1}{\varpi}L$ and hence $\mathrm{dist}([L],[\Omega_1 L]) \leq 1$, for all $[L] \in Q(\Lambda)$. □

**Lemma 5.58.** *Let $M \in \mathcal{P}_d(\mathbb{Z})$. Then $\mathrm{Id}(\mathrm{Jac}(\Lambda_{M+J_d})) = \Lambda_M$.*

*Proof.* As $\dim(Q_{M+J_d}) = d-1$, we know by Example 5.27 that the Jacobson radical of $\Lambda_{M+J_d}$ is equal to $\varpi\Lambda_M$. This is a 2-sided principal ideal in the order $\Lambda_M$, so $\mathrm{Id}(\varpi\Lambda_M) = \Lambda_M$. □

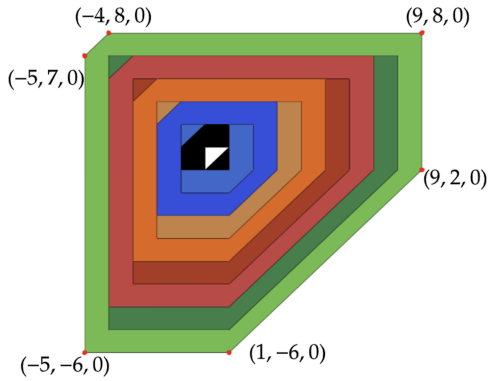



Figure 5.4: The radical idealizer process for the order $\Lambda_M$ in Example 5.59

**Example 5.59.** Consider the configuration of lattice classes $[L_{u_1}], [L_{u_2}]$ and $[L_{u_3}]$ where

$$u_1 = (0, 12, 5) \sim (-5, 7, 0), \quad u_2 = (7, 0, 6) \sim (1, -6, 0), \quad \text{and} \quad u_3 = (9, 8, 0).$$

In the notation of Section 5.2, this configuration corresponds to the matrix

$$M = \begin{pmatrix} 0 & 7 & 9 \\ 12 & 0 & 8 \\ 5 & 6 & 0 \end{pmatrix}$$

and the decreasing sequence of polytropes $(Q(\Omega_i))_{i\geq 0}$ corresponding to the radical idealizer process for the order $\Lambda_M$ is depicted in Figure 5.4. As expected, the last polytrope (in white) is indeed a simplex.

### 5.3.4 Ball Orders

In this section, we define and study a first subfamily of the bolytrope orders, namely closed orders whose set of invariant lattices is a ball in $\mathcal{B}_d^0(K)$.

**Definition 5.60.** A *ball order* in $K^{d\times d}$ is an order of the form $\mathbb{B}_r([L]) := \mathrm{PZ}(\mathrm{B}_r([L]))$, where $L$ is a lattice in $K^d$ and $r \in \mathbb{Z}_{\geq 0}$

**Theorem 5.61.** *Let $L$ be a lattice in $K^d$, $(e_1, \ldots, e_d)$ a basis of $L$ and $r$ a non-negative integer. Then, with respect to $(e_1, \ldots, e_d)$, we have*

$$\mathbb{B}_r([L]) = \{X \in \Lambda_{rJ_d} \colon X_{11} \equiv \ldots \equiv X_{dd} \bmod \varpi^r\}.$$

*Moreover, $Q(\mathbb{B}_r([L])) = \mathrm{B}_r([L])$ and the ball $\mathrm{B}_r([L])$ is PZ-closed.*

*Proof.* Put $\Lambda = \{X \in \Lambda_{rJ_d} \colon X_{11} \equiv \ldots \equiv X_{dd} \bmod \varpi^r\}$ and $\Gamma = \mathrm{End}_{\mathcal{O}_K}(L) = \Lambda(0)$. It follows from the definition of $\Lambda$ that $\varpi^r\Gamma \subseteq \Lambda$. If $L'$ is another lattice such that $\varpi^r L \subseteq L' \subseteq L$, then $\varpi^r\Gamma L' \subseteq \varpi^r\Gamma L = \varpi^r L \subseteq L'$, which yields $\varpi^r\Gamma \subseteq \mathbb{B}_r([L])$. Now the lattice classes at distance at most $r$ from $[L]$ can be described as submodules of $V_r = L/\varpi^r L$. In particular, the image $\overline{\mathbb{B}_r([L])}$ of $\mathbb{B}_r([L])$ in the endomorphism ring $\mathrm{End}_{\mathcal{O}_K}(V_r) \cong (\mathcal{O}_K/\mathfrak{m}_K^r)^{d\times d}$ is equal to the collection of all endomorphisms stabilizing every submodule of $V_r$. This ensures that

$$\overline{\mathbb{B}_r([L])} = (\mathcal{O}_K/\mathfrak{m}_K^r)\,\mathrm{Id}_d = \overline{\Lambda}.$$

As both orders $\mathbb{B}_r([L])$ and $\Lambda$ contain the kernel $\varpi^r\Gamma$ of the projection $\Gamma \to \mathrm{End}_{\mathcal{O}_K}(V_r)$, we conclude that $\Lambda = \mathbb{B}_r([L]) = \mathrm{PZ}(\mathrm{B}_r([L]))$. We now show that $Q(\mathbb{B}_r([L])) = \mathrm{B}_r([L])$. To this end, let $[L'] \in Q(\Lambda)$. Then $[\Gamma L'] \in Q(\Gamma) = \{[L]\}$. Replacing $L'$ by some homothetic lattice we hence may assume that $\Gamma L' = L$. But $\varpi^r\Gamma \subseteq \Lambda \subseteq \mathrm{End}_{\mathcal{O}_K}(L')$ so $\varpi^r\Gamma L' = \varpi^r L \subseteq L'$ so $[L'] \in \mathrm{B}_r([L])$. □

**Remark 5.62.** (Radical idealizer chain of ball orders) Let $r$ be a positive integer. Then the Jacobson radical of the ball order $\mathbb{B}_r([L]) = \mathrm{PZ}(\mathrm{B}_r([L]))$ is $\mathrm{Jac}(\mathbb{B}_r([L])) = \varpi\mathbb{B}_{r-1}([L])$, because $\varpi\mathbb{B}_{r-1}([L])$ is a pro-nilpotent ideal of $\mathbb{B}_r([L])$ with simple quotient $\mathbb{B}_r([L])/\varpi\mathbb{B}_{r-1}([L])$ isomorphic to $\mathcal{O}_K/\mathfrak{m}_K$. Now $\varpi\mathbb{B}_{r-1}([L])$ is a principal 2-sided ideal of $\mathbb{B}_{r-1}([L])$ so

$$\mathrm{Id}(\mathrm{Jac}(\mathbb{B}_r([L]))) = \mathrm{Id}(\varpi\mathbb{B}_{r-1}([L])) = \mathbb{B}_{r-1}([L])$$

and the radical idealizer chain for ball orders is thus

$$\mathbb{B}_r([L]) \subset \mathbb{B}_{r-1}([L]) \subset \ldots \subset \mathbb{B}_1([L]) \subset \mathbb{B}_0([L]) = \mathrm{End}_{\mathcal{O}_K}(L).$$

The corresponding chain of PZ-closed subsets of $\mathcal{B}_d^0(K)$ is

$$\mathrm{B}_r([L]) \supset \mathbb{B}_{r-1}([L]) \supset \ldots \supset \mathbb{B}_1([L]) \supset \mathbb{B}_0([L]) = \{[L]\}.$$

The knowledge of the radical idealizer chain of ball orders allows to prove strong properties of ball orders, like the following.

**Proposition 5.63.** *Let $r$ be a positive integer and $\Lambda$ a closed order in $K^{d\times d}$ such that $\mathrm{Id}(\mathrm{Jac}(\Lambda)) = \mathbb{B}_{r-1}([L])$. Then one has $\mathbb{B}_r([L]) \subseteq \Lambda \subseteq \mathbb{B}_{r-1}([L])$.*

*Proof.* It follows from the hypotheses and the combination of Lemma 5.57 with Theorem 5.61 that $\mathbb{B}_{r-1}([L]) \subseteq Q(\Lambda) \subseteq \mathrm{B}_r([L])$. The orders being closed, Remark 5.62 yields that $\mathbb{B}_r([L]) \subseteq \Lambda \subseteq \mathbb{B}_{r-1}([L])$. □

**Definition 5.64.** Let $r$ be a non-negative integer and $L$ a lattice in $K^d$. A *star configuration* $\star_r([L])$ with center $[L]$ and radius $r$ is a set

$$\star_r([L]) = \{[L_1], \ldots, [L_d], [L_{d+1}]\}$$

such that the following hold:

(1) $\varpi^r L \subseteq L_1, \ldots, L_{d+1} \subseteq L$,

(2) for each $i \in \{1, \ldots, d+1\}$, one has $L_i/\varpi^r L \cong \mathcal{O}_K/\mathfrak{m}_K^r$,

(3) for each $i \in \{1, \ldots, d+1\}$, one has $L = \sum_{j \neq i} L_j$.

When $r = 1$, i.e. when $\mathcal{O}_K/\mathfrak{m}_K^r$ is a field, the 1-dimensional free $\mathcal{O}_K/\mathfrak{m}_K^r$-modules $L_i/\varpi^r L$ of $L/\varpi^r L$ form a projective basis. In this sense, Definition 5.64 generalizes the definition of a projective basis to modules over rings.

**Theorem 5.65.** *Let $r$ be a non-negative integer and let $L$ be a lattice in $K^d$. Let, moreover, $\star_r([L])$ denote a star configuration with center $[L]$ and radius $r$. Then one has*

$$\mathbb{B}_r([L]) = \mathrm{PZ}(\star_r([L])) \quad \text{and} \quad \deg(\mathbb{B}_r([L])) \leq d.$$

*Proof.* Write $\Lambda := \mathrm{PZ}(\star_r([L]))$ and $\star_r([L]) =: \{[L_1], \ldots, [L_{d+1}]\}$. Since $\star_r([L])$ has radius $r$, we have that $\star_r([L]) \subseteq \mathrm{B}_r([L])$, so $\Lambda \supseteq \mathbb{B}_r([L])$. We now claim that $\Lambda$ stabilizes all lattices $L'$ with $\varpi^r L \subseteq L' \subseteq L$. To this end, write $\overline{L} = L/\varpi^r L$ and use the bar notation for the submodules of $\overline{L}$. For $1 \leq i \leq d$ let $e_i \in L_i$ be such that $\overline{\mathcal{O}_K e_i} = \overline{L_i}$. Since $L_1 + \ldots + L_d = L$, the set $\{\overline{e_1}, \ldots, \overline{e_d}\}$ is a basis of the free module $\overline{L}$. So there are $a_i \in \mathcal{O}_K$ such that $\overline{\mathcal{O}_K \sum_{i=1}^d a_i e_i} = \overline{L_{d+1}}$. Since $\star_r([L])$ is a star configuration, all $a_i$'s are units, so, replacing $e_i$ by $a_i e_i$, we assume, without loss of generality, that $L_{d+1} = \mathcal{O}_K(e_1 + \ldots + e_d) + \varpi^r L$. Since each $L_i$ is $\Lambda$-stable, the image of $\Lambda$ in $\mathrm{End}(\overline{L}) \cong (\mathcal{O}_K/\mathfrak{m}_K^r)^{d \times d}$ consists of scalar matrices and so all submodules of $\overline{L}$ are stable. This yields the claim and so $\mathbb{B}_r([L]) = \mathrm{PZ}(\star_r([L]))$. The order $\mathbb{B}_r([L])$ has degree at most $d$, because a star configuration has cardinality $d+1$. □

The following remark shows that ball orders in $K^{d \times d}$ can have degree (in the sense of Definition 5.52) smaller than $d$.

**Remark 5.66.** The degree of $\mathbb{B}_r([\mathcal{O}_K^4])$ is at most 3, because $\mathbb{B}_r([\mathcal{O}_K^4])$ is equal to the Plesken-Zassenhaus order of the following lattices (where the columns of the matrices are the basis elements):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & \varpi^r & 0 \\ 0 & 0 & 0 & \varpi^r \end{pmatrix}, \begin{pmatrix} \varpi^r & 0 & 0 & 0 \\ 0 & \varpi^r & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & \varpi^r & 0 \\ 1 & 0 & 0 & \varpi^r \end{pmatrix}, \text{ and } \begin{pmatrix} \varpi^r & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \varpi^r & 0 \\ 0 & 1 & 0 & \varpi^r \end{pmatrix}.$$

Via change of coordinates, one obtains that any ball order in $K^{4 \times 4}$ has degree at most 3.

### 5.3.5 Bolytrope Orders

Let $M \in \mathcal{P}_d(\mathbb{Z})$. Recall, from Definition 5.42, that the bolytrope $\mathrm{B}_r(M)$ is defined to be $\mathrm{B}_r(Q(\Lambda_M))$.

**Definition 5.67.** A *bolytrope order* is an order of the form $\mathbb{B}_r(M) := \mathrm{PZ}(\mathrm{B}_r(M))$, where $M$ is an element of $\mathcal{P}_d(\mathbb{Z})$ and $r$ is a non-negative integer.

Until the end of the present section, fix $M \in \mathcal{P}_d(\mathbb{Z})$ and an apartment $\mathcal{A}$ containing $Q(\Lambda_M)$. Let, moreover, $r$ be a non-negative integer. Then, by Lemma 5.43, we have that $\mathrm{B}_r(M) \cap \mathcal{A} = Q(\Lambda_{M+rJ_d})$, in particular $\mathbb{B}_r(M) \subseteq \Lambda_{M+rJ_d}$. Put

$$\Lambda_r(M) = \{X \in \Lambda_{M+rJ_d} \colon X_{11} \equiv \ldots \equiv X_{dd} \bmod \varpi^r\}.$$

We will show that $\Lambda_r(M) = \mathbb{B}_r(M)$ and $Q(\Lambda_r(M)) = \mathrm{B}_r(M)$ is PZ-closed; cf. Theorem 5.72.

**Lemma 5.68.** *Let $[L]$ be a lattice class in $Q(\Lambda_M)$. Then $\Lambda_r(M) = \Lambda_{M+rJ_d} \cap \mathbb{B}_r([L])$ and $\Lambda_r(M)$ is a closed order.*

*Proof.* Let $(e_1, \ldots, e_d)$ be a basis of $L$ that is also a frame basis defining the apartment $\mathcal{A}$. Then, with respect to this basis, $[L] = [\mathcal{O}_K^d]$ and thus $\Lambda_M \subseteq \mathrm{End}_{\mathcal{O}_K}(L) = \mathcal{O}_K^{d\times d} = \Lambda_{0^{d\times d}}$. It follows in particular that $M$ has non-negative entries. The explicit description of the ball order in Theorem 5.61 allows to deduce that $\Lambda_r(M) = \Lambda_{M+rJ_d} \cap \mathbb{B}_r([L])$. Since $\Lambda_{M+rJ_d}$ and $\mathbb{B}_r([L])$ are closed orders, then so is $\Lambda_r(M)$. □

**Lemma 5.69.** *One has $\mathrm{B}_r(M) \subseteq Q(\Lambda_r(M))$ and $\Lambda_r(M) \subseteq \mathbb{B}_r(M)$.*

*Proof.* We first show that $\mathrm{B}_r(M) \subseteq Q(\Lambda_r(M))$. To that end, let $[L'] \in \mathrm{B}_r(M)$ and let $[L] \in Q(\Lambda_M)$ be such that $\mathrm{dist}([L'], [L]) \leq r$. Then the combination of Remark 5.62 and Lemma 5.68 yields

$$[L'] \in \mathrm{B}_r([L]) = Q(\mathbb{B}_r([L])) \subseteq Q(\Lambda_r(M)).$$

To conclude, the inclusion $\mathrm{B}_r(M) \subseteq Q(\Lambda_r(M))$ implies that $\Lambda_r(M) \subseteq \mathbb{B}_r(M)$. □

To prove that $\mathbb{B}_r(M) = \Lambda_r(M)$ we use the radical idealizer chain of $\Lambda_r(M)$, which we describe in the following remark.

**Remark 5.70.** Assume that $r \geq 1$. Then, similarly to what is done in Remark 5.62, one sees that $\mathrm{Jac}(\Lambda_r(M)) = \varpi\Lambda_{r-1}(M)$ is a 2-sided principal ideal of $\Lambda_{r-1}(M)$ and hence $\mathrm{Id}(\mathrm{Jac}(\Lambda_r(M))) = \Lambda_{r-1}(M)$.

**Lemma 5.71.** *One has $Q(\Lambda_r(M)) = \mathrm{B}_r(M)$.*

*Proof.* Lemma 5.69 shows that $\mathrm{B}_r(M) \subseteq Q(\Lambda_r(M))$. For the opposite inclusion, we rely on Remark 5.70 to proceed by induction on $r$. Assume first that $r = 0$. Then $Q(\Lambda_0(M)) = Q(\Lambda_M) = \mathbb{B}_0(M)$ and so we are done. Now assume that $r > 0$ and that $Q(\Lambda_{r-1}(M)) = \mathbb{B}_{r-1}(M)$. The fact that $\Lambda_{r-1}(M) = \mathrm{Id}(\mathrm{Jac}(\Lambda_r(M)))$ together with Lemma 5.57 then yields that

$$Q(\Lambda_r(M)) \subseteq \mathbb{B}_1(Q(\Lambda_{r-1}(M))) = \mathbb{B}_1(\mathbb{B}_{r-1}(M)) \subseteq \mathrm{B}_r(M).$$

This concludes the proof. □

The following is the main result of this section.

**Theorem 5.72.** *The following hold:*

$$\Lambda_r(M) = \mathbb{B}_r(M) \quad \text{and} \quad Q(\mathbb{B}_r(M))) = \mathrm{B}_r(M).$$

*In particular, bolytrope orders are closed, and bolytropes are* PZ*-closed.*

*Proof.* As a consequence of Lemma 5.68, both $\Lambda_r(M)$ and $\mathbb{B}_r(M)$ are closed orders. We are now done thanks to Lemma 5.71. □

**Corollary 5.73.** *The beginning of the radical idealizer chain for bolytrope orders is*

$$\mathbb{B}_r(M) \subset \mathbb{B}_{r-1}(M) \subset \ldots \subset \mathbb{B}_1(M) \subset \mathbb{B}_0(M) = \Lambda_M.$$

*The first $r+1$ elements in the corresponding chain of* PZ*-closed subsets of $\mathcal{B}_d^0(K)$ are*

$$\mathrm{B}_r(M) \supset \mathrm{B}_{r-1}(M) \supset \ldots \supset \mathrm{B}_1(M) \supset Q(\Lambda_M).$$

Note that $\Lambda_M$ is the first term in the radical idealizer process that is a graduated order. The polytrope $Q(\Lambda_M)$ is hence canonically determined by the bolytrope $\mathrm{B}_r(M)$ and called the *central polytrope* of $\mathrm{B}_r(M)$.

In analogy with ball orders, we obtain the following stronger property of bolytrope orders.

**Corollary 5.74.** *Assume that $r \geq 1$ and let $\Lambda$ be a closed order in $K^{d\times d}$ such that we have $\mathrm{Id}(\mathrm{Jac}(\Lambda)) = \mathbb{B}_{r-1}(M)$. Then $\mathbb{B}_r(M) \subseteq \Lambda \subseteq \mathbb{B}_{r-1}(M)$.*

*Proof.* Analogous to the proof of Proposition 5.63. □

The following theorem states that any bolytrope (and hence also any bolytrop order) can be determined by at most $d+1$ lattice classes. That is any bolytrope order is of degree at most $d$, in the sense of Definition 5.52.

**Theorem 5.75.** *Let $[P_1], \ldots, [P_d]$ be the distinct classes of projective $\Lambda_{M+rJ_d}$-lattices. Then there is a lattice class $[L_{d+1}] \in \mathrm{B}_r(M)$, such that*

$$\mathbb{B}_r(M) = \mathrm{PZ}([P_1], \ldots, [P_d], [L_{d+1}]).$$

*Moreover, the degree of $\mathbb{B}_r(M)$ is at most $d$.*

*Proof.* As a consequence of Remark 5.46, we have that $\Lambda_{M+rJ_d} = \mathrm{PZ}([P_1], \ldots, [P_d])$. In particular, for any lattice class $[L_{d+1}] \in \mathrm{B}_r(M)$, Lemma 5.68 and Theorem 5.72 imply that

$$\mathbb{B}_r(M) \subseteq \mathrm{PZ}([P_1], \ldots, [P_d], [L_{d+1}]) \subseteq \Lambda_{M+rJ_d}.$$

To construct $L_{d+1}$ such that the inclusion $\mathbb{B}_r(M) \supseteq \mathrm{PZ}([P_1], \ldots, [P_d], [L_{d+1}])$ holds, choose $[L] \in Q(\Lambda_M)$ and a lattice basis $(e_1, \ldots, e_d)$ of $L$ that is also a frame basis for some apartment

containing $Q(\Lambda_{M+rJ_d})$. Define $L_{d+1} := \mathcal{O}_K(e_1 + \ldots + e_d) + \varpi^r L$ and, for each $i = 1, \ldots, d$, put $L_i := \mathcal{O}_K e_i + \varpi^r L \in Q(\Lambda_{M+rJ_d})$. Then $\{[L_1], \ldots, [L_d], [L_{d+1}]\}$ is a star configuration with center $[L]$ and radius $r$. By Theorem 5.65, we thus have

$$\mathrm{PZ}([L_1], \ldots, [L_d], [L_{d+1}]) = \mathbb{B}_r([L]),$$

which, together with Lemma 5.68 and Theorem 5.72, implies that the order

$$\mathrm{PZ}([P_1], \ldots, [P_d], [L_{d+1}])$$

is contained in

$$\Lambda_{M+rJ_d} \cap \mathbb{B}_r([L]) = \Lambda_r(M) = \mathbb{B}_r(M).$$

□

### 5.3.6 When the building is a tree

Throughout this section, assume that $d = 2$. Then the building $\mathcal{B}_2(K)$ is an infinite tree. Apartments correspond to infinite paths in the tree and the bounded convex subsets of $\mathcal{B}_2(K)$ are the bounded subtrees. For more on this and other trees, see for instance [148].

The following is the main result of this section. It extends [164, Theorem 2] beyond the case of finite residue fields.

**Theorem 5.76.** *Let $\Lambda$ be a closed order in $K^{2\times 2}$. Then there are $r, m \in \mathbb{Z}_{\geq 0}$ such that*

$$\Lambda = \mathbb{B}_r\left(\begin{pmatrix} 0 & m \\ 0 & 0 \end{pmatrix}\right) = \{X \in \mathcal{O}_K^{2\times 2} \colon X_{12} \in \mathfrak{m}_K^{m+r}, X_{21} \in \mathfrak{m}_K^r, X_{11} \equiv X_{22} \bmod \varpi^r\}.$$

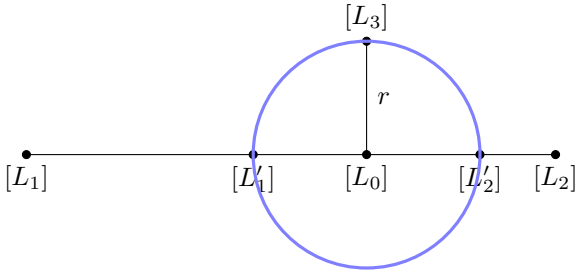
*Proof.* Put $R := \max\{\mathrm{dist}([L], [L']) \colon [L], [L'] \in Q(\Lambda)\}$ and let $[L_1], [L_2] \in Q(\Lambda)$ be such that $R = \mathrm{dist}([L_1], [L_2])$. Then the convex hull

$$\mathcal{L} = Q(\mathrm{PZ}([L_1], [L_2])) \subseteq Q(\Lambda)$$

is a line segment and is hence contained in an apartment $\mathcal{A}$. Define

$$r := \max\{\mathrm{dist}([L], \mathcal{L}) | [L] \in Q(\Lambda)\}$$

and let $[L_3] \in Q(\Lambda)$ be such that $r = \mathrm{dist}([L_3], \mathcal{L})$. Let, moreover, $[L_0] \in \mathcal{L}$ denote the unique lattice class in $\mathcal{L}$ satisfying $\mathrm{dist}([L_3], [L_0]) = r$.



Now choose a frame basis $(e_1, e_2)$ for $\mathcal{A}$ such that, with respect to this basis, there exists an integer $m$ such that $[L_1] = [L_{(0,r)}]$ and $[L_2] = [L_{(m+r,0)}]$. It follows from the definition of $R$ that

$$R + 1 = |\mathcal{L}| = m + 2r + 1.$$

With respect to the chosen basis, note now that $\mathcal{L} = Q(\Lambda_{M+rJ_2})$ and hence

$$\Lambda \subseteq \Lambda_{M+rJ_2}.$$

Moreover, if $[L_1']$ and $[L_2'] \in \mathcal{L}$ are the two lattice classes at distance $r$ from $[L_0]$ and such that $\mathrm{dist}([L_1'], [L_2']) = 2r$, then the set $\{[L_3], [L_1'], [L_2']\}$ is a star configuration with radius $r$ and center $[L_0]$. As a consequence of the definition of $\mathcal{L}$, such lattice classes $[L_1'], [L_2']$ exist and thus Theorem 5.65 ensures that

$$\Lambda \subseteq \mathbb{B}_r([L_0]).$$

We have proven that $\Lambda \subseteq \mathbb{B}_r([L_0]) \cap \Lambda_{M+rJ_2}$ and so $\Lambda \subseteq \mathbb{B}_r(M)$, thanks to Lemma 5.68. As $Q(\Lambda) \subseteq \mathrm{B}_r(M) = Q(\mathbb{B}_r(M))$, we obtain $\Lambda = \mathbb{B}_r(M)$ as stated in the theorem. □

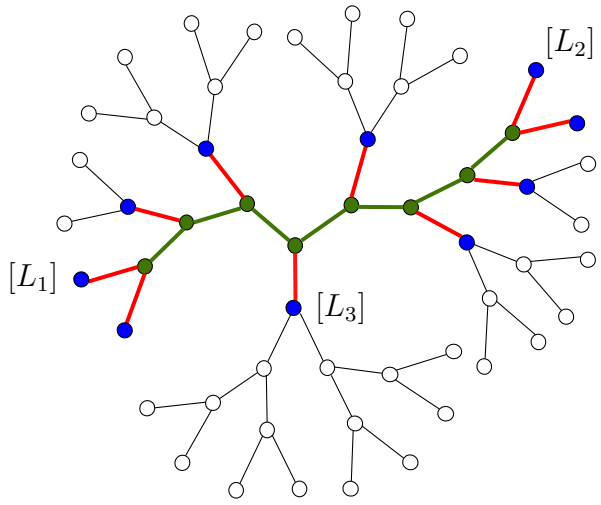



Figure 5.5: The bolytrope $\mathbb{B}_1(Q) = \mathbb{B}_1\left(\begin{pmatrix} 0 & 7 \\ 0 & 0 \end{pmatrix}\right)$ in the Bruhat Tits tree of $\mathrm{SL}_2(\mathbb{Q}_2)$. The green segment is the central polytrope $Q := Q\left(\begin{pmatrix} 0 & 7 \\ 0 & 0 \end{pmatrix}\right) = \{[L_{(i,0)} \colon 0 \leq i \leq 7\}$. The set $\mathcal{L} = Q\left(\begin{pmatrix} 0 & 8 \\ 1 & 0 \end{pmatrix}\right)$ is the convex hull of $[L_1] = [L_{(0,1)}]$ and $[L_2] = [L_{(8,0)}]$. The blue vertices are the points at distance 1 from $Q$. The PZ-order of the lattice classes $[L_1], [L_2]$ and $[L_3]$ is the same as the PZ-order of all the colored vertices.

**Corollary 5.77.** *The* PZ*-closed subset of* $\mathcal{B}_2^0(K)$ *are precisely the bolytropes.*

**Corollary 5.78.** *The degree of a closed order* $\Lambda$ *in* $K^{2\times 2}$ *is* 0, 1, *or* 2. *Orders of degree* 0 *are the maximal orders, whereas the closed orders of degree* 1 *are precisely the graduated non-maximal orders. All non-graduated closed orders in* $K^{2\times 2}$ *have degree* 2.

**Remark 5.79.** Theorem 5.75 implies [164, Theorems 1 and 8]. To see this, note that, by taking $[L_1], [L_2], [L_3]$ as in the proof of Theorem 5.76, we get that $\Lambda = \mathrm{PZ}([L_1], [L_2], [L_3])$.

## 5.4 Conclusion

In conclusion, graduated orders are Plessken-Zassenhaus rings of configurations of lattice classes that are contained in one apartment. The structure of these orders is very tightly linked to the tropical geometry and combinatorics of their set of invariant lattices which is a polytrope in the tropical torus. Bolytrope orders are a generalization of Graduated orders. They arise as the Plessken-Zassenhaus rings of the Minkowski sum of a polytrope and a ball in the Bruhat-Tits building $\mathcal{B}_d(K)$.

# Chapter 6

# Non-archimedean Schur representations of $\mathrm{GL}(n, \mathcal{O}_K)$ and invariant lattices

This chapter is based on joint work [51] with Antonio Lerario.

## 6.1 Introduction

Let $K$ be a discretely valued non-archimedean field and $\mathcal{O}_K$ its valuation ring. Given a vector space $V$ of dimension $n$ over $K$ (identified with $K^n$ through the choice of a basis) and a partition $\lambda$ of an integer $d$, in this chapter we address the problem of determining lattices in the Schur module $S_\lambda(V)$ which are invariant under the action of $\mathrm{GL}(n, \mathcal{O}_K)$ given by the representation

$$\rho_{n,\lambda} : \mathrm{GL}(n, \mathcal{O}_K) \to \mathrm{GL}(S_\lambda(V)). \tag{6.1}$$

Our motivation for this problem has a probabilistic origin. In fact, in the case $K = \mathbb{Q}_p$, $\mathcal{O}_K = \mathbb{Z}_p$ and $\lambda = (d)$, we have $S_\lambda(V) \simeq \mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$, the space of homogenous polynomials in $n$ variables of degree $d$, with the action of $\mathrm{GL}(n, \mathbb{Z}_p)$ by linear change of variables. The problem of determining lattices in $\mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ which are invariant under this action is equivalent to the problem of determining Gaussian probability distributions on $\mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ (in the sense of Evans, see Section 1.1.4, Chapter 3 and [68, 61, 63, 64, 65, 66]) which are invariant under this action and has recently become of particular interest in the context of the emerging field of "Probabilistic Algebraic Geometry" (see Chapter 3). We will discuss this connection with more details in Section 6.1.2 (see also Section 6.2). The next theorem is our main result.

**Theorem 6.1.** *Let $p$ be the residue characteristic of $K$ and $\lambda$ be a partition of $d \geq 1$. If none the hook lengths of $\lambda$ is divisible by $p$, then there exists a unique lattice in $S_\lambda(V)$ (up to scaling) which is invariant under the action of $\mathrm{GL}(n, \mathcal{O}_K)$.*

**Remark 6.2.** The condition that the partition $\lambda$ is $p$-core is necessary. For a counter example, see Example 6.12.

Recall that the hook length of a box in the Young diagram of a partition $\lambda$ is obtained by adding 1 to the number of boxes below and to the right of the box in question. If a prime $\ell$ does not divide any of the hook lengths of $\lambda$, the partition is said to be *$\ell$-core* (by convention every partition is 0-core). In particular if $p > d$ then $\lambda$ is $p$-core and the hypotheses of Theorem 6.1 are satisfied.

### 6.1.1 Group actions on the Bruhat-Tits Buildings.

Our main result can be restated more precisely (see Theorem 6.11) using the language of *Buildings*. These are combinatorial and geometric structures that generalize certain aspects of Riemannian symmetric spaces, see [94, 141]. These structures were introduced to better understand reductive algebraic groups via their action on such structures which are of interest in geometric group theory [52, 55, 136, 137].

In our context, observe that lattices in $S_\lambda(V)$ (up to homothety) are points in the *Bruhat-Tits building* $\mathcal{B}_{n,\lambda}$ for $\mathrm{PGL}(S_\lambda(V))$. This is an infinite simplicial complex whose 0-simplices are the homothety classes of lattices in $S_\lambda(V)$, see Section 1.2. The action of $\mathrm{GL}(n, \mathcal{O}_K)$ on $S_\lambda(V)$ induces an action on the building $\mathcal{B}_{n,\lambda}$ i.e. we have a group homomorphism

$$\rho^{\mathcal{B}}_{n,\lambda} \colon \mathrm{GL}(n, \mathcal{O}_K) \to \mathrm{Aut}(\mathcal{B}_{n,\lambda}),$$

sending $\mathrm{GL}(n, \mathcal{O}_K)$ to the group of automorphisms of the building. With this notation, Theorem 6.11 below states that if $\lambda$ is $\mathrm{char}(K)$-core, the set of fixed points of this action is a non-empty finite convex set in the building $\mathcal{B}_{n,\lambda}$, reduced to one point if $\lambda$ is $p$-core, where $p$ is the residue characteristic of $K$.

### 6.1.2 Probabilistic Algebraic Geometry.

In the last years there has been an increasing interest into the statistical behaviour of algebraic sets over non-algebraically closed fields. When the notion of "generic" is no longer available, one seeks for a "random" study of the objects of interest. For instance, once a probability distribution is put on the space $K[x_1, \ldots, x_n]_{(d)}$ of homogeneous polynomials, one can study expected properties of their zero sets in projective space $\mathbb{P}^{n-1}_K$ (e.g. the expected number of solutions of systems of random equations), see for instance [5, 16, 28, 45, 46, 48, 50, 54, 76, 74, 75, 102, 105, 109, 110, 128, 143, 149, 150, 151].

The appropriate choice of a norm on $K^n$ (using a scalar product when $K = \mathbb{R}$, a hermitian product when $K = \mathbb{C}$ or a non-archimedean norm when $K = \mathbb{Q}_p$) induces a metric structure on $\mathbb{P}^{n-1}_K$ and the group $\mathrm{Iso}(K^n) \subset \mathrm{GL}(n, K)$ of linear norm-preserving transformations, acts by isometries on the projective space. In this context it is natural to put a probability distribution on the space of polynomials $K[x_1, \ldots, x_n]_{(d)}$ which is invariant under the action of $\mathrm{Iso}(K^n)$ induced by change of variables – there should be no preferred points or directions

in projective space. An interesting problem is therefore to find (and possibly classify) all the probability distributions on $K[x_1, \dots, x_n]_{(d)}$ having this property.

When $K = \mathbb{R}, \mathbb{C}$, and in the case of nondegenerate *Gaussian* distributions, this problem is equivalent to the problem of finding scalar products (hermitian structures in the complex case) on $K[x_1, \dots, x_n]_{(d)}$ which are preserved by orthogonal (unitary in the complex case) change of variables, and it was solved by Kostlan [101] using representation theory.

If $K = \mathbb{C}$, then $\mathrm{Iso}(\mathbb{C}^n) \simeq U(n, \mathbb{C})$ and, since the change of variables representation

$$\rho_{n,d}^{\mathbb{C}} : U(n, \mathbb{C}) \to \mathrm{GL}(\mathbb{C}[x_1, \dots, x_n]_{(d)})$$

is irreducible, by Schur's Lemma there is a unique such hermitian structure (up to multiples).

If $K = \mathbb{R}$, then $\mathrm{Iso}(\mathbb{R}^n) \simeq \mathrm{O}(n, \mathbb{R})$ and the representation

$$\rho_{n,d}^{\mathbb{R}} : \mathrm{O}(n, \mathbb{R}) \to \mathrm{GL}(\mathbb{R}[x_1, \dots, x_n]_{(d)})$$

*is not* irreducible. Its irreducible summands are spaces of spherical harmonics and there is a whole $\lfloor \frac{d}{2} \rfloor$-dimensional family of scalar products having the desired property, see [101].

If $K = \mathbb{Q}_p$, then $\mathrm{Iso}(\mathbb{Q}_p^n) \simeq \mathrm{GL}(n, \mathbb{Z}_p)$ (see [69, Theorem 2.4]). Moreover, Evans [68, 61, 63, 64, 65, 66] introduced a notion of *Gaussian* distribution on a $p$-adic vector space, which essentially corresponds to a choice of lattice in the vector space (in the same way as real Gaussian structures corresponds to scalar products, see Section 1.1.4 and Section 6.2). Using this correspondence, in this context the above problem can be formulated as: find all lattices in $\mathbb{Q}_p[x_1, \dots, x_n]_{(d)}$ which are invariant under the action of $\mathrm{GL}(n, \mathbb{Z}_p)$ by change of variables.

Our Theorem 6.1 implies that, if $\lambda$ is $p$-core (in particular if $p > d$), there is only one such lattice (up to scaling) and therefore only one probability distribution on $\mathbb{Q}_p[x_1, \dots, x_n]_{(d)}$ with the required property (up to scaling). See Section 6.2 for more details.

Here again the irreducibility of $\rho_{n,d}^{\mathbb{Q}_p} : \mathrm{GL}(n, \mathbb{Z}_p) \to \mathrm{GL}(\mathbb{Q}_p[x_1, \dots, x_n]_{(d)})$ (Theorem 6.10 below) plays a role for the uniqueness, but in a different way (compared to the complex case). More generally, we have the following result, proved in subsection 6.4.3.

**Corollary 6.3.** *Suppose $K$ is a non-archimedean local field and assume that $\lambda$ is $\mathrm{char}(K)$-core. Then there are only finitely many Gaussian distributions on $S_\lambda(V)$ (up to scaling) that are invariant under the action of $\mathrm{GL}(n, \mathcal{O}_K)$ through $\rho_{n,\lambda}$. Moreover, if $\lambda$ is also $p$-core, where $p$ is the residue characteristic of $K$, there is only one such measure (up to scaling).*

This chapter is organized as follows. In Section 6.2, we explain the probabilistic motivation behind the work presented in this chapter. We collect some background and preliminary results in Section 6.3, and prove our main results in Section 6.4. Finally, we discuss some open questions and final remarks in Section 6.5.

## 6.2 Probabilistic motivation

We recall in this section some standard notions from probability, to help the reader to put our results into context. The reader may skip this section and come back to it later.

### 6.2.1 Gaussian distributions on real vector spaces

We start by recalling the notion of Gaussian distributions on a real vector space. First, the *standard Gaussian distribution* on $\mathbb{R}^m$ is the probability measure $\gamma_m$ on $\mathbb{R}^m$ defined for every Borel set $U \subseteq \mathbb{R}^m$ by

$$\gamma_m(U) := \frac{1}{(2\pi)^{\frac{m}{2}}} \int_U e^{-\frac{\|x\|^2}{2}} dx,$$

where $\|x\| := \sqrt{|x_1|^2 + \cdots + |x_n|^2}$ denotes the Euclidean norm of $x = (x_1, \ldots, x_m)$. In practice, if $\{e_1, \ldots, e_m\}$ is an orthonormal basis for the Euclidean norm and $\xi_1, \ldots, \xi_m$ are independent, standard Gaussians (i.e. $\mathbb{P}(\xi_j \leq t) = \gamma_1(-\infty, t)$ for every $j = 1, \ldots, m$), sampling from the standard Gaussian distribution on $\mathbb{R}^m$ is equivalent to picking an element $\xi \in \mathbb{R}^m$ at random by writing it as a random linear combination

$$\xi = \xi_1 e_1 + \cdots + \xi_m e_m.$$

More generally, a nondegenerate, centered, Gaussian distribution on a real finite dimensional vector space $V$ is given by assigning a surjective linear map $T : \mathbb{R}^m \to V$ and defining $\mathbb{P}(W) := \gamma_m(T^{-1}(W))$ for every Borel set $W \subseteq V$. (This measure is denoted by $T_{\#}\gamma_m$ and called the *push-forward* measure.) Writing $\mathbb{R}^m = \mathrm{Ker}(T) \oplus \mathrm{Ker}(T)^{\perp}$, one gets a linear isomorphism $V \simeq \mathrm{Ker}(T)^{\perp}$ and an induced scalar product on $V$. As above, if $\{v_1, \ldots, v_n\}$ denotes an orthonormal basis for $V$, this construction is equivalent to define a random element in $V$ by $\xi_1 v_1 + \cdots + \xi_n v_n$, with the $\xi_j$'s standard, independent Gaussians. Viceversa, given a scalar product on $V$ and an orthonormal basis $\{v_1, \ldots, v_n\}$ for it, putting $\xi := \xi_1 v_1 + \cdots + \xi_n v_n$, where the $\xi_i$ are standard, independent Gaussians, defines a random variables with values in $V$ and the induced probability distribution is nondegenerate, centered and Gaussian (as a map $T : \mathbb{R}^n \to V$ in this case we could simply take $T(e_i) := v_i$).

From this we see that the theory of nondegenerate, centered, Gaussian distributions on $V$ is equivalent to the theory of nondegenerate positive definite quadratic forms on it (i.e. scalar products). In particular, once a representation $\rho : G \to \mathrm{GL}(V)$ is given, asking for the nondegenerate, centered Gaussian probability distributions $\gamma$ on $V$ which are $\rho$-invariant (i.e. such that $\rho(g)_{\#}\gamma = \gamma$ for all $g \in G$) is equivalent to ask for the scalar products $\langle \cdot, \cdot \rangle$ on $V$ which are $\rho$-invariant (i.e. such that $\langle v_1, v_2 \rangle = \langle \rho(g)v_1, \rho(g)v_2 \rangle$ for all $v_1, v_2 \in V$ and $g \in G$).

In particular, if the representation $\rho$ is irreducible, Schur's Lemma implies that there is only one such invariant scalar product (up to multiples) and, consequently, only one invariant Gaussian distribution (up to scaling). If the representation *is not* irreducible, this is no longer true as it happens for the case of real polynomials.

### 6.2.2 Invariant Gaussian distributions on the space of real and complex polynomials

Let now $\rho^{\mathbb{C}}_{n,d} : U(n,\mathbb{C}) \to \mathrm{GL}(\mathbb{C}[x_1,\dots,x_n]_{(d)})$ be the representation given by change of variables:

$$\rho^{\mathbb{C}}_{n,d}(g)(P) := P \circ g^{-1}.$$

This representation is complex irreducible and, consequently, there is only one $\rho^{\mathbb{C}}_{n,d}$-invariant hermitian structure (up to multiples) on $\mathbb{C}[x_1,\dots,x_n]_{(d)}$. It is called the *Bombieri-Weyl* hermitian structure. A hermitian orthonormal basis for it is given by the monomials

$$\left\{ \left( \frac{d!}{\alpha_1! \cdots \alpha_n!} \right)^{1/2} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \right\}_{\alpha_1+\cdots+\alpha_n=d}. \tag{6.2}$$

The Bombieri-Weyl scalar product and the corresponding Gaussian distribution have been widely used, see [45, 46, 74, 75, 76, 102, 109, 110, 128, 143, 149, 150, 151].

Notice that the basis in (6.2) is real and, since $\mathrm{O}(n,\mathbb{R}) \subset U(n,\mathbb{R})$, the Bombieri–Weyl scalar product restricts to a scalar product on $\mathbb{R}[x_1,\dots,x_n]_{(d)}$ which is invariant under $\mathrm{O}(n,\mathbb{R})$. In other words, denoting by $\rho^{\mathbb{R}}_{n,d} : \mathrm{O}(n,\mathbb{R}) \to \mathrm{GL}(\mathbb{R}[x_1,\dots,x_n]_{(d)})$ the representation by change of variables, the Bombieri-Weyl scalar product is $\rho^{\mathbb{R}}_{n,d}$-invariant. However, since $\rho^{\mathbb{R}}_{n,d}$ *is not* real irreducible, there are invariant scalar products which are not multiples of this one. The classification of such scalar products has been done by Kostlan in [101] using the theory of spherical harmonics, as follows. Denoting by $\mathcal{H}_{n,\ell} \subset \mathbb{R}[x_1,\dots,x_n]_{(\ell)}$ the space of harmonic polynomials, we have a decomposition

$$\mathbb{R}[x_1,\dots,x_n]_{(d)} \simeq \bigoplus_{d-\ell \text{ even}} \|x\|^{d-\ell} \cdot \mathcal{H}_{n,\ell}.$$

The spaces $\|x\|^{d-\ell} \cdot \mathcal{H}_{n,\ell}$ are precisely the irreducible summands of $\rho^{\mathbb{R}}_{n,d}$ and are isomorphic to the spaces of spherical harmonics. Schur's Lemma implies that there is only one invariant scalar product on each of them, up to multiples, and an invariant scalar product on $\mathbb{R}[x_1,\dots,x_n]_{(d)}$ is obtained by scaling these scalar products separately. For instance, the $L^2(S^{n-1})$ scalar product is $\rho^{\mathbb{R}}_{n,d}$-invariant but is not a multiple of the Bombieri-Weyl one (see [73]).

#### 6.2.2.1 Gaussian distributions on $p$-adic spaces

Evans [61, 66, 63, 64, 65, 68] (see also Section 1.1.4) introduced the notion of Gaussian distribution on a $p$-adic vector space, as follows. To start with, one denotes by $\zeta_1$ the uniform probability measure on the compact topological group $\mathbb{Z}_p$ and by $\zeta_m := \zeta_1 \times \cdots \times \zeta_1$ the product measure on $\mathbb{Z}_p^m$ (these are just the normalized Haar measures). The measure $\zeta_m$ is called the *standard $p$-adic Gaussian measure*. Then, if $V \simeq \mathbb{Q}_p^n$ is a $p$-adic vector space, a $p$-adic Gaussian measure is defined by assigning a surjective linear map $T : \mathbb{Q}_p^m \to V$ and considering the pushforward measure $T_\# \zeta_m$ (notice the analogy with the real construction).

Here the theory of nondegenerate $p$-adic Gaussian measures on $V \simeq \mathbb{Q}_p^n$ is equivalent to the theory of full dimensional lattices in $V$: the image of $\mathbb{Z}_p^m$ under $T$ is a lattice $L := T(\mathbb{Z}_p^m)$ in $V$ and every lattice arise in this way. This lattice is the support of the measure $T_{\#}\zeta_m$. Given a lattice $L \subset V \simeq \mathbb{Q}_p^n$, one defines a $p$-adic Gaussian distribution fixing a basis $\{v_1, \ldots, v_n\}$ for $L$ and setting $T(e_j) = v_j$, as above (here $\{e_1, \ldots, e_n\}$ is the standard basis). A random element from this distribution is obtained by setting $\xi := \xi_1 v_1 + \cdots + \xi_n v_n$, where the $\xi_j$'s are independent uniform random variables on $\mathbb{Z}_p$. Gaussian distributions on a discretely valued non-archimedean field $K$ are constructed similarly.

Given a representation $\rho : G \to \mathrm{GL}(V)$, it is then natural to ask for lattices which are invariant under this representation (or equivalently invariant under the action of the ring and $\mathcal{O}_K$-module $H_\rho \coloneqq \mathrm{span}_R(\mathrm{im}(\rho)) \subset \mathrm{End}_K(V)$); they correspond to $\rho$-invariant Gaussian distributions. Unlike the real and complex setting, in the non-archimedean case, even if $\rho$ is irreducible, Schur's Lemma cannot be used to conclude uniqueness in this context. That is because non-degenerate Gaussian measures are in a one-to-one correspondence with lattices (instead of positive non-degenerate quadratic forms as in the real case).

### 6.2.3 Invariant Gaussian distributions on the space of $p$-adic polynomials

Evans [67] defines a probability distribution on the space of polynomials considering the random polynomial

$$\zeta(y) := \zeta_0 + \zeta_1 \binom{y}{1} + \cdots + \zeta_d \binom{y}{d}, \tag{6.3}$$

where $\zeta_0, \ldots, \zeta_d$ are independent and uniformly distributed in $\mathbb{Z}_p$. Since the aforementioned seminal work of Evans a couple of decades ago, probabilistic problems over non-archimedean local fields have been gaining interest in the recent years [5, 16, 28, 48, 50, 54]. In many of these problems, it is important that the probability measure is invariant under certain symmetries. Homogenizing the above polynomial, one gets a probability distribution on $\mathbb{Q}_p[x, y]_{(d)}$ which *is not* invariant under the action of $\mathrm{GL}(2, \mathbb{Z}_p)$ by change of variables. Here $\mathrm{GL}(n, \mathbb{Z}_p)$ can be seen as the group of isometries of projective space $\mathbb{P}_{\mathbb{Q}}^{n-1}$ (see [105]), and it is natural to ask for a probability distribution on $\mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ for which there are no preferred points or direction for the zero sets of polynomials in projective space, as we did for the real case.

In [105] the authors proposed an alternative model, defining a random polynomial $\zeta \in \mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ as

$$\zeta(x) := \sum_{|\alpha| = \alpha_1 + \cdots + \alpha_n = d} \zeta_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \tag{6.4}$$

where $\{\xi_\alpha\}_{|\alpha|=d}$ is a family of indenendent random variables, each of them uniformly distributed in $\mathbb{Z}_p$. This Gaussian distribution corresponds to the lattice $L \subset \mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ spanned by the standard monomial basis.

The probability distribution induced by Equation (6.4) on $\mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ is invariant under the action of $\mathrm{GL}(n, \mathbb{Z}_p)$ by change of variables (see [105]). It is natural to ask whether there are other Gaussian distributions on $\mathbb{Q}_p[x_1, \ldots, x_n]_{(d)}$ which have this property. Our Theorem 6.1 implies that, when $p$ does not divide $d$, Equation (6.4) is the only distribution with this property (up to scaling). In fact, this corresponds to the case $\lambda = (d)$ with the Young diagram

$$\lambda = \begin{array}{|c|c|c|c|} \hline 1 & 2 & \cdots & d \\ \hline \end{array}$$

More generally, the same result is true for a discretely valued non-archimedean field $K$: Corollary 6.3 states that if $\lambda$ is $p$-core, with $p$ the residue characteristic, then there is a unique Gaussian measure on the space $S_\lambda(V) \cong K[x_1, \ldots, x_n]_{(d)}$ of homogeneous polynomials of degree $d$ in $n$ variables that is invariant under linear change of variables in $\mathrm{GL}(n, \mathcal{O}_K)$.

## 6.3 Supporting results

In this section, we introduce some notation, collect the necessary background. We also prove some preliminary results we will need in Section 7.4.

### 6.3.1 The group $\mathrm{GL}(n, \mathcal{O}_K)$

As before, let $K$ be a non-archimedean discretely valued field with normalized valuation $\mathrm{val}\colon K^\times \twoheadrightarrow \mathbb{Z}$. We denote by $\mathcal{O}_K$ its valuation ring and we fix a uniformizer $\varpi$ of $K$. We denote by $k$ the residue field $\mathcal{O}_K/\varpi\mathcal{O}_K$, and $\ell, p$ the respective characteristics [1] of the fields $K$ and $k$. The non-archimedean valuation val induces an ultrametric absolute value $|\cdot|$ on $K$ as follows:

$$|x| \coloneqq \begin{cases} p^{-\mathrm{val}(x)}, & \text{if } p > 0 \\ e^{-\mathrm{val}(x)}, & \text{otherwise} \end{cases}.$$

There is a natural notion of non-archimedean orthogonality (see [65, Section 3] for more details) for which the analogue of the group of orthogonal $n \times n$ matrices $\mathrm{O}(n, \mathbb{R})$ in the real setting is the group

$$\mathrm{GL}(n, \mathcal{O}_K) \coloneqq \left\{ g \in \mathrm{GL}(n, K)\colon g, g^{-1} \in \mathcal{O}_K^{n \times n} \right\},$$

see [69, Theorem 2.4]; this group is a totally disconnected compact topological group.

[1] When $\ell > 0$ we necessarily have $p = \ell$.

### 6.3.2 The Schur functor

Let $V$ be an $n$-dimensional $K$-vector space[2] spanned by a basis $x_1, \ldots, x_n$. There exists a natural linear right action of the group $\mathrm{GL}(n, \mathcal{O}_K)$ on $V$ as follows

$$x_i \cdot g = \sum_{j=1}^{n} g_{ij}\ x_j, \quad g = (g_{ij})_{1 \le i,j \le n} \in \mathrm{GL}(n, \mathcal{O}_K).$$

This clearly defines a faithful representation $\rho_n \colon \mathrm{GL}(n, \mathcal{O}_K) \to \mathrm{GL}(V)$. Let $d \ge 1$ and $\lambda \vdash d$ a partition of $d$. The *Schur functor* $S_\lambda$ maps the representation $\rho_n$ to the representation $\rho_{n,\lambda}$ acting on the *Weyl module* $S_\lambda(V)$ defined as

$$S_\lambda(V) \coloneqq c_\lambda \cdot V^{\otimes d},$$

where $c_\lambda$ is the *Young symmetrizer*. For a detailed construction of the Weyl module $S_\lambda(V)$ we refer the reader to [72, Section 6.1], [71, Chapter 8] or [83, Chapter 6].

Theorem 1 in section 8.1 of [71] gives a basis of $S_\lambda(V)$ whose elements are indexed by the Young tableaux $T$ obtained by filling the Young diagram of $\lambda$ with entries in $\{1, \ldots, d\}$. Choosing this basis as the *standard basis* of $S_\lambda(V)$ we get a group homomorphism

$$\rho_{n,\lambda} \colon \mathrm{GL}(n, K) \to \mathrm{GL}(N, K),$$

with $N \coloneqq \dim_K(S_\lambda(V))$.

### 6.3.3 Lattices

*Lattices* in $S_\lambda(V)$ are full rank $\mathcal{O}_K$-submodules of $S_\lambda(V)$. The representation $\rho_{n,\lambda}$ defines a natural action of $\mathrm{GL}(n, \mathcal{O}_K)$ on lattices of $S_\lambda(V)$ as follows:

$$g \cdot L \coloneqq \{g \cdot x \colon x \in L\}, \qquad \text{as an } \mathcal{O}_K\text{-module.}$$

In this chapter, we are interested in determining the lattices that are invariant under this action i.e. lattices $L$ such that

$$g \cdot L = L, \quad \text{for all } g \in \mathrm{GL}(n, \mathcal{O}_K).$$

(Obviously, if $L$ is $\rho_{n,\lambda}$-invariant, then $aL$ is $\rho_{n,\lambda}$-invariant for any $a \in K^\times$. So, to be more precise, we are interested in $\rho_{n,\lambda}$-invariant *homothety classes* of lattices.)

A central object for our study is the $\mathcal{O}_K$-submodule of $\mathrm{End}_K(S_\lambda(V))$ generated by the image of $\rho_{n,\lambda}$, which we denote by $H_{n,\lambda}$ i.e.

$$H_{n,\lambda} = \mathrm{span}_R(\mathrm{im}(\rho_{n,\lambda})).$$

[2]Note that same construction can be carried out more generally when $V$ is an $\mathcal{O}_K$-module.

The $\mathcal{O}_K$-module $H_{n,\lambda}$ is also a subring of $\mathrm{End}_K(S_\lambda(V))$. Given a lattice $L$ in $S_\lambda(V)$, the action of $\rho_{n,\lambda}$ defines a new lattice $H_{n,\lambda} \cdot L$ defined as follows

$$H_{n,\lambda} \cdot L = \sum_{f \in H_{n,\lambda}} f \cdot L.$$

The above sum (6.3.3) is finite, since we can take the sum over an $\mathcal{O}_K$-basis of $H_{n,\lambda}$. The module $H_{n,\lambda}$ is a torsion-free over the discrete valuation ring $\mathcal{O}_K$, so it is free and has a finite $\mathcal{O}_K$-basis. Notice also that, since $\mathrm{id} \in H_{n,\lambda}$ we always have $L \subset H_{n,\lambda} \cdot L$.

### 6.3.4 Some auxiliary results from representation theory

Given the standard basis of $S_\lambda(V)$, the representation $\rho_{n,\lambda}$ maps a matrix $g \in \mathrm{GL}(n, K)$ to a matrix $\rho_{n,\lambda}(g) \in \mathrm{GL}(N, K)$ wih $N = \dim_K(S_\lambda(V))$; the entries of the matrix $\rho_{n,\lambda}(g)$ are homogeneous polynomials of degree $d$ in the entries of $g$.

**Example 6.4.** Suppose that $n = d = 2$ and

$$\lambda = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline \end{array}$$

Then the space $S_\lambda(V)$ is then the second symmetric power of $V = K \cdot x_1 \ \oplus \ K \cdot x_2$ i.e. the space of homogeneous degree 2 polynomials in 2 variables

$$S_\lambda(V) = K \cdot x_1^{\otimes 2} \ \oplus \ K \cdot (x_1 \otimes x_2 + x_2 \otimes x_1) \ \oplus \ K \cdot x_2^{\otimes 2}.$$

The representation $\rho_{2,\lambda}$ can then be described in a matrix form as follows

$$\rho_{2,\lambda} \colon \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a^2 & 2ac & c^2 \\ ab & ad + bc & cd \\ c^2 & 2bd & d^2 \end{bmatrix}.$$

See [122, Chapter 11] for similar explicit The ring $H_{2,\lambda}$, in matrix form, given by

$$H_{2,\lambda} \coloneqq \left\{ X \in R^{3\times 3} \colon X_{12}, X_{32} \in 2\mathcal{O}_K \right\}.$$

Let us recall the following result from representation theory of algebraic groups.

**Theorem 6.5.** *Let $F$ be a field and $V$ a vector space over $F$. Suppose $\lambda$ is $\mathrm{char}(F)$-core. Then the Weyl module $S_\lambda(V)$ is absolutely irreducible as a representation of the group $\mathrm{GL}(n, F)$. Moreover, the map*

$$\begin{aligned} \Phi \colon S_\lambda(V) \otimes S_\lambda(V)^* &\to \mathcal{O}(\mathrm{GL}(n, K)) \\ v \otimes \beta &\mapsto (g \mapsto \langle \beta, \ \rho_{n,\lambda}(g) \ v \rangle), \end{aligned} \tag{6.5}$$

*where $S_\lambda(V)^*$ is the dual space of $S_\lambda(V)$ and $\mathcal{O}(\mathrm{GL}(n, K))$ is the Hopf algebra of regular functions on $\mathrm{GL}(n, K)$, is injective.*

*Proof.* When $\mathrm{char}(F) = 0$ the result can be immediately deduced from the *algebraic Peter-Weyl theorem*, which states that for a reductive algebraic group $G$ we have

$$\mathcal{O}(G) = \bigoplus_{(\pi,W) \text{ irrep of } G} W \otimes W^*,$$

where the sum is over the irreducible representations $\pi$ of $G$. For a reference, see [160, Theorem 27.3.9] or [80, Theorem 12.1.4].

The positive characteristic case is known to the experts, but we could not find a precise reference. Hence for completeness, we provide a concise proof. Let $G := \mathrm{GL}(n, F)$ and assume that $\lambda$ is $\mathrm{char}(F)$-core and denote by $W_\lambda$ the Specht module over $F$ associated to $\lambda$ (this is a representation of the symmetric group $S_d$). Since $\lambda$ is a $\mathrm{char}(F)$-core partition, the Specht module $W_\lambda$ is an absolutely irreducible and projective $S_d$-module. Then by the Schur-Weyl duality the Weyl module $S_\lambda(V) \cong \mathrm{Hom}_{S_d}(W_\lambda, V^{\otimes d})$ is an absolutely irreducible representation of $G$. So the division algebra $D = \mathrm{End}_G(S_\lambda(V))$ of intertwining operators is trivial i.e. $D \cong F$ and using the Jacobson density theorem we deduce that the $F$-linear map

$$F[G] \twoheadrightarrow \mathrm{End}_D(S_\lambda(V)) = \mathrm{End}_F(S_\lambda(V)), \quad g \mapsto \rho_{n,\lambda}(g)$$

is surjective, where $F[G]$ is the free group algebra of $G$ over $F$. This implies that if $\alpha \in \mathrm{End}_F(S_\lambda(V))^*$ is a linear form with $\alpha(\rho_{n,\lambda}(g)) = 0$ for all $g \in G$, then $\alpha(f) = 0$ for $f \in \mathrm{End}_F(S_\lambda(V)$ i.e. $\alpha = 0$. We then conclude that the map $\Phi$ is indeed injective. □

**Lemma 6.6.** *Let $A$ be a principal ideal domain with infinitely many units and $F = \mathrm{Frac}(A)$ its field of fractions. Let $\varphi \in \mathcal{O}(\mathrm{GL}(n, F))$ be a polynomial function on $\mathrm{GL}(n, K)$. If $\varphi(g) = 0$ for all $g \in \mathrm{GL}(n, A)$, then $\varphi = 0$.*

*Proof.* Suppose that $\varphi(h) = 0$ for any $h \in \mathrm{GL}(n, A)$ and that $\varphi \neq 0$. Then there exists $g \in \mathrm{GL}(n, F)$ with $\varphi(g) = 0$. Since $A$ is a principal ideal domain, we can write the Smith normal form $g = ug'v$ of $g$ where $u, v \in \mathrm{GL}(n, A)$ and $g'$ is diagonal. So, by replacing $\varphi$ with the polynomial $x \mapsto \varphi(uxv)$, we may assume, without loss of generality, that $g$ is a diagonal matrix. Since $\varphi(g) \neq 0$ and $g$ is diagonal, the restriction of $\varphi$ to the space of diagonal matrices is a nonzero polynomial

$$\phi(z_1, \ldots, z_n) = \varphi(\mathrm{diag}(z_1, \ldots, z_n)) = \sum_{\nu \in \mathbf{N}^n} c_\nu \, z_1^{\nu_1} \ldots z_n^{\nu_n}.$$

Since $\varphi$ vanishes on $\mathrm{GL}(n, A)$ we deduce that

$$\varphi(\mathrm{diag}(u_1, \ldots, u_n)) = \phi(u_1, \ldots, u_n) = 0, \quad \text{for } u_1, \ldots, u_n \in A^\times.$$

Since $A$ has infinitely many units, we deduce that $\phi = 0$ which is a contradiction. So, as desired, we conclude that $\varphi = 0$. □

**Remark 6.7.** The statement of Lemma 6.6 clearly fails when $A$ is a domain with finitely many units. For example if $\mathcal{O}_K = \mathbb{Z}$, the function $g \mapsto (\det(g) - 1)(\det(g) + 1)$ vanishes on $\mathrm{GL}(n, \mathbb{Z})$ but not on $\mathrm{GL}(n, \mathbb{Q})$.

**Proposition 6.8.** *If $\lambda$ is $\ell$-core, the $\mathcal{O}_K$-module $H_{n,\lambda}$ spans $\mathrm{End}_K(S_\lambda(V))$ over $K$.*

*Proof.* Suppose that $\lambda$ is $\ell$-core and assume that $H_{n,\lambda}$ does not span $\mathrm{End}_K(S_\lambda(V))$. Then there exists a non-zero linear form $\alpha \in \mathrm{End}_K(S_\lambda(V))^*$ such that

$$\alpha(f) = 0, \quad \text{for all } f \in H_{n,\lambda}. \tag{6.6}$$

Equivalently, we then have $\alpha(\rho_{n,\lambda}(g)) = 0$ for all $g \in \mathrm{GL}(n, \mathcal{O}_K)$. Then, by virtue of Lemma 6.6 we deduce that

$$\alpha(\rho_{n,\lambda}(g)) = 0, \quad \text{for all } g \in \mathrm{GL}(n, K).$$

Here we are extending $\rho_{n,\lambda}$ in the obvious way and using the fact that $\mathrm{GL}(n, \mathcal{O}_K)$ is an open set in $\mathrm{GL}(n, K)$. Identifying $\mathrm{End}_K(S_\lambda(V))^*$ with $S_\lambda(V) \otimes S_\lambda(V)^*$ in the canonical way, (6.6) can then be rewritten as $\Phi(\alpha) = 0$. But, since $\Phi$ is injective by virtue of Theorem 6.5, we deduce that $\alpha = 0$ which is a contradiction. Hence $H_{n,\lambda}$ spans $\mathrm{End}(S_\lambda(V))$ over $K$. □

**Remark 6.9.** In the language of Chapter 6, when $H_{n,\lambda}$ spans $\mathrm{End}_K(S_\lambda(V))$, we say that $H_{n,\lambda}$ is an *order* in $\mathrm{End}_K(S_\lambda(V))$ i.e. a $\mathcal{O}_K$-module of full rank that is also a ring.

## 6.4 Proofs of main results

### 6.4.1 Irreducibility of Schur representations

In this section we prove the following result, giving a sufficient condition for the irreducibility of $(\rho_{n,\lambda}, S_\lambda(V))$. Recall that, a representation $\rho : G \to \mathrm{GL}(V)$ is said to be irreducible if it has no proper invariant subspace.

**Theorem 6.10.** *If $\lambda$ is $\ell$-core, then the representation $(\rho_{n,\lambda}, S_\lambda(V))$ of the group $\mathrm{GL}(n, \mathcal{O}_K)$ is irreducible.*

*Proof.* Suppose that $\lambda$ is $\ell$-core and assume that there is a proper subspace $W \subset S_\lambda(V)$ which is $\rho_{n,\lambda}$–invariant; that is

$$\rho_{n,\lambda}(g)(W) \subset W, \quad \text{for all } g \in \mathrm{GL}(n, \mathcal{O}_K).$$

But, since $W$ is a proper subspace of $S_\lambda(V)$, the linear space

$$\mathrm{Stab}(W) \coloneqq \{f \in \mathrm{End}_K(S_\lambda(V)) \colon f(W) \subset W\},$$

is a proper subspace of $\mathrm{End}_K(S_\lambda(V))$. Since $\rho_{n,\lambda}(g) \in \mathrm{Stab}(W)$ we deduce that $H_{n,\lambda} \subset \mathrm{Stab}(W)$. By virtue of Proposition 6.8, this is a contradiction. So we deduce that $\rho_{n,\lambda}$ is an irreducible representation of $\mathrm{GL}(n, \mathcal{O}_K)$. □

### 6.4.2 Invariant lattices: the fixed point set in the Bruhat-Tits building

Recall that the action of $\mathrm{GL}(n,\mathcal{O}_K)$ on $S_\lambda(V)$ induces an action on the building $\mathcal{B}_{n,\lambda}$ i.e. we have a group homomorphism

$$\rho^{\mathcal{B}}_{n,\lambda}\colon \mathrm{GL}(n,\mathcal{O}_K)\to \mathrm{Aut}(\mathcal{B}_{n,\lambda}),$$

sending $\mathrm{GL}(n,\mathcal{O}_K)$ to the group of automorphisms of the building $\mathcal{B}_{n,\lambda}$. We denote by $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ the set of 0-simplices of $\mathcal{B}_{n,\lambda}$ that are fixed under $\rho^{\mathcal{B}}_{n,\lambda}$ i.e.

$$\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda}) \coloneqq \{[L]\colon [L]=\rho^{\mathcal{B}}_{n,\lambda}(g)([L]) \text{ for all } g\in \mathrm{GL}(n,\mathcal{O}_K)\}.$$

Passing to the quotient modulo $\varpi$, the representation $\rho_{n,\lambda}$ of $\mathrm{GL}(n,\mathcal{O}_K)$ naturally induces the representation

$$\overline{\rho_{n,\lambda}}\colon \mathrm{GL}(n,k)\to \mathrm{GL}(S_\lambda(L_0/\varpi L_0)),$$

where $L_0 \coloneqq \mathcal{O}_K\cdot x_1\oplus\cdots\oplus\mathcal{O}_K\cdot x_n$. The following result is our main theorem.

**Theorem 6.11.** *Suppose that $\lambda$ is $\ell$-core. Then the set $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ is a non-empty finite convex set in the building $\mathcal{B}_{n,\lambda}$ (in the sense of Section 1.2 and Chapter 5). Moreover, if $\lambda$ is $p$-core, the set $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ is reduced to the one point*

$$\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda}) = \{[S_\lambda(L_0)]\},$$

*with $L_0 \coloneqq \mathcal{O}_K\cdot x_1\oplus\cdots\oplus\mathcal{O}_K\cdot x_n$.*

*Proof.* The lattice $\Lambda_0 = S_\lambda(L_0)$ is always $\rho_{n,\lambda}$-invariant so $[\Lambda_0]\in\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$. If $\Lambda_1,\Lambda_2$ are two $\rho_{n,\lambda}$-invariant lattices then $\Lambda_1+\Lambda_2$ and $\Lambda_1\cap\Lambda_2$ are also $\rho_{n,\lambda}$-invariant so we deduce that $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ is convex in $\mathcal{B}_{n,\lambda}$.

Now suppose that $\lambda$ is $\ell$-core. Then, by virtue of Proposition 6.8, the $\mathcal{O}_K$-module $H_{n,\lambda}$ is an order in $\mathrm{End}_K(S_\lambda(V))$. Hence there exists an integer $r>0$ such that $\mathrm{id}+\varpi^r\,\mathrm{End}_{\mathcal{O}_K}(L_0)\subset H_{n,\lambda}$. By virtue of Theorem 5.61, the set of points in $\mathcal{B}_{n,\lambda}$ that are invariant under the ball order $\mathrm{id}+\varpi^r\,\mathrm{End}_{\mathcal{O}_K}(\Lambda_0)$ is exactly the ball $\mathrm{B}([\Lambda_0],r)$ of center $\Lambda_0$ and radius $r$ in the building $\mathcal{B}_{n,\lambda}$ which is finite, so $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})\subset\mathrm{B}([\Lambda_0],r)$ is finite.

Now assume further that $\lambda$ is $p$-core and assume that there is a neighbor $[\Lambda_1]$ of $[\Lambda_0]$ in $\mathcal{B}_{n,\lambda}$; that is

$$\varpi\Lambda_0\subsetneq\Lambda_1\subsetneq\Lambda_0,$$

such that $[\Lambda_1]\in\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$. Then $W=\Lambda_1/\varpi\Lambda_0$ is a proper $\overline{\rho_{n,\lambda}}$-invariant subspace of $\Lambda_0/\varpi\Lambda_0 = S_\lambda(L_0/\varpi L_0)$. But since $\lambda$ is $p$-core, by virtue of Theorem 6.5, the representation $\overline{\rho_{n,\lambda}}$ is irreducible so this is a contradiction. We then deduce that no neighbour of $[\Lambda_0]$ in $\mathcal{B}_{n,\lambda}$ is $\rho_{n,\lambda}$-invariant. Finally, since $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ is convex in $\mathcal{B}_{n,\lambda}$, we conclude that $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda}) = \{[\Lambda_0]\}$. □

### 6.4.3 Proof of Corollary 6.3

Let $\mathbb{P}$ be a $\mathrm{GL}(n, \mathcal{O}_K)$-invariant Gaussian measure on $S_\lambda(V)$ (in the sense of Section 1.1.4). Let $\Lambda := \mathrm{supp}(\mathbb{P})$ be the $\mathcal{O}_K$-submodule of $S_\lambda(V)$ that is the support of $\mathbb{P}$ (see [65, Section 4]). Then since $\mathbb{P}$ is $\mathrm{GL}(n, \mathcal{O}_K)$-invariant, the module $\Lambda$ is $\mathrm{GL}(n, \mathcal{O}_K)$-invariant hence also $H_{n,\lambda}$-invariant. Assume that $\lambda$ is $p$-core, then by virtue of Proposition 6.8 $H_{n,\lambda}$ is an order. Hence either $\Lambda = 0$ or $\Lambda = H_{n,\lambda} \cdot \Lambda$ has full rank in $S_\lambda(V)$. Assuming $\Lambda \neq 0$, i.e. the measure $\mathbb{P}$ is not the Dirac measure at 0, we deduce that $\Lambda$ is a $\mathrm{GL}(n, \mathcal{O}_K)$-invariant lattice in $S_\lambda(V)$. Then using Theorem 6.11 we deduce that $[\Lambda] = [\Lambda_0]$ which finishes the proof. □

## 6.5 Concluding remarks and open questions

### 6.5.1 An example of unbounded $\mathrm{Fix}(\rho_{n,\lambda})$

In the case where $\lambda$ is not $\ell$-core (which implies that $p = \ell$), the set $\mathrm{Fix}(\rho_{n,\lambda})$ is still convex but can be unbounded in the building $\mathcal{B}_{n,\lambda}$.

**Example 6.12.** Assume that $K$ is a local field of characteristic 2 (for example the field of Laurent series $\mathbb{F}_2((\varpi))$). For any integer $m \geq 0$ let $L_m$ be the following lattice

$$L_m = \mathcal{O}_K \cdot x^2 \oplus \mathcal{O}_K \cdot y^2 \oplus \varpi^m \mathcal{O}_K \cdot xy,$$

in the space of homogeneous polynomial $S_{2,(2)}(V) \cong K[x, y]_{(2)}$. For $\alpha, \beta, \gamma \in \mathcal{O}_K$, and $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, \mathcal{O}_K)$ notice that

$$\begin{aligned}(\alpha x^2 + \beta y^2 + \gamma \varpi^m xy) \cdot g &= \alpha(ax + by)^2 + \beta(cx + dy)^2 + \gamma \varpi^m (ax + by)(cx + dy) \\ &= (\alpha a^2 + \beta b^2 + \gamma \varpi^m ac)x^2 + (\alpha c^2 + \beta d^2 + \gamma \varpi^m bd)y^2 \\ &\quad + \gamma \varpi^m (ad + bc)xy.\end{aligned}$$

So we deduce that $g \cdot L_m \subset L_m$ for any $g \in \mathrm{GL}(2, \mathcal{O}_K)$. Since the action of $\mathrm{GL}(2, \mathcal{O}_K)$ is measure preserving we deduce that $g \cdot L_m = L_m$ for any such $g$. So $\mathrm{Fix}(\rho_{2,(2)})$ is unbounded in $\mathcal{B}_{2,(2)}$ since it contains the homothety class $[L_m]$ for $m \geq 0$.

Notice also that the representation $\rho_{2,(2)}$ is not irreducible (since the space W spanned by $x^2, y^2$ is $\rho_{2,(2)}$-invariant) nor semisimple (the space $W$ has no $\mathrm{GL}(2, K)$-invariant complement).

### 6.5.2 Computing $\mathrm{Fix}(\rho_{n,\lambda})$ for small residue characteristics

In this section we focus on the mixed characteristic case when $\ell = 0$ and $\lambda$ is not $p$-core. In this case, we know that $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ is a finite convex set in $\mathcal{B}_{n,\lambda}$ and we are interested in computing lattices in $\mathrm{Fix}(\rho^{\mathcal{B}}_{n,\lambda})$ (up to scaling).

**Conjecture 6.13.** Let $N = \dim_K(S_\lambda(V))$ and suppose that $\ell = 0$. We conjecture that the order $H_{n,\lambda}$ is a *graduated order* in the sense of Chapter 5; in other words there exists a matrix $M = (m_{ij}) \in \mathbb{Z}^{N\times N}$ with

1. $m_{ii} = 0$ for any $1 \leq i \leq N$,
2. $m_{ij} \leq m_{ik} + m_{kj}$ for any $1 \leq i, j, k \leq N$,

such that the order $H_{n,\lambda}$ (in matrix form given the standard basis of $S_\lambda(V)$) is given by

$$H_{n,\lambda} = \Lambda_M \coloneqq \{X \in K^{N\times N} \colon X_{ij} \in \varpi^{m_{ij}}\mathcal{O}_K\}.$$

When $H_{n,\lambda}$ is a graduated order, the set $\mathrm{Fix}(\rho_{n,\lambda})$ is a *polytrope* that lies in one apartment of the building $\mathcal{B}_{n,\lambda}$ and can be fully determined from the matrix $M$ (see Corollary 5.13).

**Example 6.14.** Assume that $\ell = 0$ and $p = 2$ (this is the case for example whenever $K$ is a finite field extension of $\mathbb{Q}_2$). In Example 6.4, the order $H_{2,(2)}$ in matrix form is $\Lambda_M$ with

$$M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

So by virtue of Corollary 5.13 we deduce that $\mathrm{Fix}(\rho_{2,(2)})$ consists of two points

$$[S_{(2)}(\mathcal{O}_K \cdot x_1 \oplus \mathcal{O}_K \cdot x_2)] \quad \text{and} \quad [\varpi\mathcal{O}_K \cdot x_1^{\otimes 2} \ \oplus \ \mathcal{O}_K \cdot (x_1 \otimes x_2 + x_2 \otimes x_1) \ \oplus \ \varpi\mathcal{O}_K \cdot x_2^{\otimes 2}].$$

**Remark 6.15.** In the equal characteristic $\ell = p$, the set $\mathrm{Fix}(\rho_{n,\lambda})$ is either reduced to one point (when $\lambda$ is $\ell$-core) or is unbounded in $\mathcal{B}_{n,\lambda}$ (when $\lambda$ is not $\ell$-core).

### 6.5.3 Other compact group actions

In addition to the action of $\mathrm{GL}(n, \mathcal{O}_K)$ on the Bruhat-Tits building $\mathcal{B}_{n,\lambda}$, one could study the action of closed subgroups of $\mathrm{GL}(n, \mathcal{O}_K)$. The groups $\mathrm{SL}(n, \mathcal{O}_K)$, $\mathrm{SO}(n, \mathcal{O}_K)$ and the symmetric group $S_n$ are of particular interest. One might suspect that the same results proven in this chapter hold also for $\mathrm{SL}(n, \mathcal{O}_K)$. For $\mathrm{SO}(n, \mathcal{O}_K)$ the representation $(\rho_{n,\lambda}, S_\lambda(V))$ will no longer be irreducible and we suspect that it decomposes into a sum of irreducible representations in a similar way[3] it does over the real numbers for $\mathrm{SO}(n, \mathbb{R})$. The set of fixed lattices for $\mathrm{SO}(n, \mathcal{O}_K)$ will then be unbounded, as will be the case for $S_n$ (see Example 2.31). It would be interesting to describe what these convex sets look like in the building.

[3]Because a linear space being stable under a polynomial group action is a purely algebraic fact.

# Chapter 7

# Tropical invariants for binary quintics and reduction types of Picard curves

This chapter is based on joint work [49] with Paul A. Helminck and Enis Kaya. We express the reduction types of Picard curves in terms of tropical invariants associated to binary quintics. We furthermore give a general framework for tropical invariants associated to group actions on arbitrary varieties. The problem of describing reduction types of curves in terms of their associated invariants fits in this general framework by mapping the space of binary forms to symmetrized versions of the Deligne–Mumford compactification $\overline{M}_{0,n}$.

## 7.1 Introduction

Invariant theory studies quantities in geometry that are invariant under group actions. This theory sparked many developments in commutative algebra, leading to the Hilbert basis theorem and many other results. In this chapter, we study invariants of binary forms $f(x,z) = a_0x^n + a_1x^{n-1}z + \cdots + a_nz^n$ defined over an algebraically closed field $K$ of characteristic 0. The group in question is $\mathrm{GL}(2,K)$ and it acts on these binary forms through *Möbius transformations*; that is

$$f^\sigma(x,z) = f(ax+bz, cx+dz), \quad \text{for } \sigma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2,K).$$

By comparing the entries of $f$ to those of $f^\sigma$, this then also gives an action of $\mathrm{GL}(2,K)$ on $K[a_i]$. The invariants for this action are homogeneous polynomials $H$ in the $a_i$ such that $H^\sigma = \det(\sigma)^k H$ for some $k$. By Hilbert's basis theorem, these form a finitely generated subring of $K[a_i]$ called the ring of invariants. There are algorithms that can explicitly calculate the generators of this ring and a full list of generators is known for binary forms of low degrees, see [154]. These generators satisfy the pleasant property that two separable binary forms are projectively equivalent if and only if the values of the generators are projectively equivalent. This gives a strong connection between algebra on the one hand and geometry on the other.

We now turn to the non-archimedean side of this story and consider a complete non-archimedean algebraically closed field $K$ of characteristic zero with non-trivial valuation $v\colon K^* \to \mathbb{R}$. Let $f(x,z)$ be a separable binary form of degree $n$ over $K$. The zeroes of this form give a canonical metric tree on $n$ leaves by connecting the corresponding points in the Berkovich analytification of $\mathbb{P}^1$. There are only finitely many phylogenetic types for any given degree $n$ (see Figure 7.1 for the case $n = 5$), and the possible types give rise to a partition of the space of all non-archimedean binary forms of degree $n$. Since the tree is invariant under projective isomorphisms, this also partitions the space of all invariants. A natural question now arises: what are the equations for these partitions?

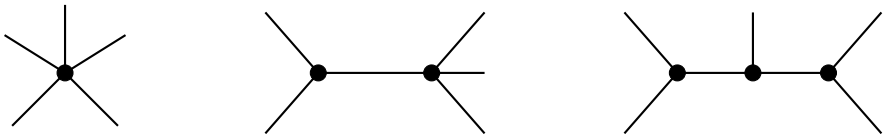



Figure 7.1: The three types of unmarked phylogenetic trees with 5 leaves

There are two instances where we know the equations: $n = 4$ and $n = 6$. For $n = 4$ there are two tree types: a trivial one with four leaves connected to a single vertex and a non-trivial one. We can distinguish between these two using the valuation of the $j$-invariant of the quartic. Namely, a quartic has trivial tree type if and only if $\operatorname{val}(j) \geq 0$, see [152, Chapter VII]. For $n = 6$, there are seven tree types and one can distinguish between them using the valuations of Igusa invariants, see [115] and [89]. Our main goal in this chapter is to fill up this gap and find the invariants for *binary quintics*, that is, for $n = 5$.

For quintics, there are three tree types; see Figure 7.1. We wish to distinguish between these tree types using the valuations of suitable invariants. To that end, we start with a set of generators $I_4, I_8, I_{12}, I_{18}$ for the ring of invariants together with the discriminant $\Delta$. The valuations of these invariants are not sufficient to determine the tree type of the quintic (see Example 7.27), so we introduce a new invariant $H$, giving the set $S = \{I_4, I_8, I_{12}, I_{18}, \Delta, H\}$. In our first theorem, we show that the valuations of these invariants determine the tree type of a quintic. We call the valuations of these invariants the *tropical invariants* of the quintic. For technical reasons, we assume for the remainder of this section that the residue characteristic $p$ of $K$ is not equal to $2, 3, 11$. Note however that Remark 7.28 explains how to deal with the case where $p = 11$.

**Theorem 7.1** (**T**ree types of binary quintics)**.** *Let $f$ be a separable binary quintic over $K$. Then, the tree type of $f$ is determined by the tropical invariants as follows:*

*(I) The tree is of Type* I *if and only if* $8\operatorname{val}(I) - \deg(I)\operatorname{val}(\Delta) \geq 0$ *for all* $I \in S$.

*(II) The tree is of Type* II *if and only if* $\operatorname{val}(\Delta) - 2\operatorname{val}(I_4) > 0$ *or* $9\operatorname{val}(\Delta) - 4\operatorname{val}(I_{18}) > 0$, *and* $12\operatorname{val}(I) - \deg(I)\operatorname{val}(H) \geq 0$ *for all* $I \in S$.

*(III) The tree is of Type* III *if and only if* $\operatorname{val}(\Delta) - 2\operatorname{val}(I_4) > 0$ *and* $\operatorname{val}(H) - 3\operatorname{val}(I_4) > 0$.

Finding the equations for $n = 4$ and $n = 6$ is partially motivated by applications to reduction types of hyperelliptic curves. For $n = 5$, the motivation comes from Picard curves $X$, which are smooth plane quartics of the form

$$y^3\ell(x, z) = q(x, z),$$

where $\deg(\ell(x, z)) = 1$ and $\deg(q(x, z)) = 4$. We are interested in obtaining the *minimal skeleton* of the Berkovich analytification of such a curve, which codifies the different possible semistable models for $X$. The results in [87] show that this skeleton can be recovered from the *marked tree type* of the quintic $f(x, z) = \ell(x, z) \cdot q(x, z)$. More precisely, this quintic has five distinct roots, giving a metric tree with five leaves, and the root of $\ell(x, z)$ gives the marking. We can assume by a projective transformation that this marked point is $\infty$. There are exactly five marked tree types, see Figure 7.2, giving rise to five reduction types of Picard curves. In terms of invariant theory, the natural object to consider here is the binary $(4, 1)$-form $(q(x, z), \ell(x, z))$. These binary $(4, 1)$-forms similarly have a finitely generated ring of invariants and in our second theorem we show that we can find a set of invariants for $(4, 1)$-forms that distinguish between the five marked tree types. This set consists of the set $S$ from Theorem 7.1 together with a new set $S' = \{j_2, j_3, j_5, j_6, j_9\}$ of $(4, 1)$-invariants. As above, we call the valuations of these invariants the *tropical invariants* of the $(4, 1)$-form.

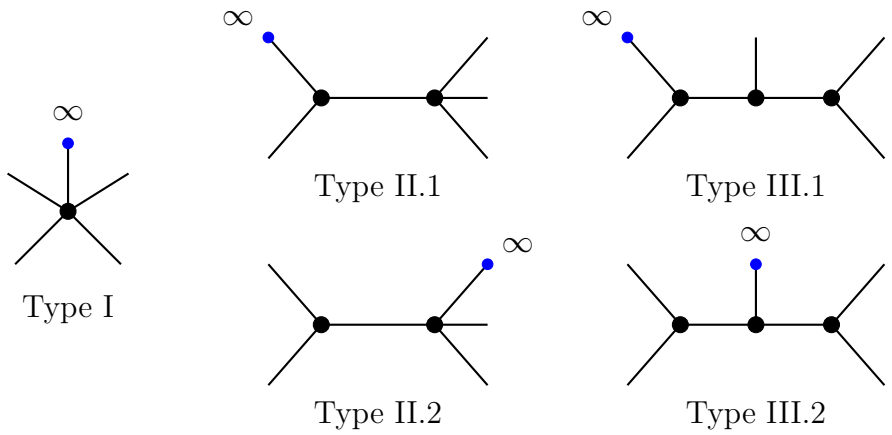



Figure 7.2: Tree types of binary $(4, 1)$-forms

**Theorem 7.2 (Tree types of $(4, 1)$-forms).** *Let $(q, \ell)$ be a $(4, 1)$-form over $K$ such that the associated binary quintic $f = q \cdot \ell$ is separable. The tree type of $(q, \ell)$ is determined by the tropical invariants as follows:*

*(I) If $f$ has tree Type* I*, then $(q, \ell)$ also has Type* I.

*(II) If $f$ has tree Type* II*, then $(q, \ell)$ has Type* II.1 *(resp. Type* II.2*) if and only if the quantity $5\,\mathrm{val}(j_2) - 2\,\mathrm{val}(j_5)$ is strictly positive (resp. zero).*

*(III) If $f$ has tree Type* III*, then $(q, \ell)$ has Type* III.1 *(resp. Type* III.2*) if and only if the quantity $5\,\mathrm{val}(j_2) - 2\,\mathrm{val}(j_5)$ is strictly positive (resp. zero).*

To determine the skeleton of a Picard curve, we also need to know the associated weights and edge lengths. The weights of the skeleton are completely determined by Theorem 7.2, but the lengths are not. In our third theorem, we give formulas for the edge lengths of a $(4, 1)$-form in terms of its tropical invariants. For trees of Type II and Type III.2, we are able to give these in terms of invariants of quintics. For trees of Type III.1, we express the marked edge lengths in terms of $(4, 1)$-invariants. This difference is quite natural, since there is a natural symmetry on trees of Type III.2.

**Theorem 7.3** (**Edge lengths**)**.** *The non-trivial edge lengths of the trees in Theorem 7.2 are given by the tropical invariants as follows:*

*(I) If $f$ has Type* I*, then there are no non-trivial edges.*

*(II) If $f$ has Type* II*, then there is only one edge $e_1$. Its length, both in the cases of unmarked and marked trees, is given by*

$$L(e_1) = \max\left(\frac{1}{2}(\mathrm{val}(\Delta) - 2\,\mathrm{val}(I_4)),\ \frac{1}{3}(2\,\mathrm{val}(\Delta) - \mathrm{val}(I_{18}))\right).$$

*(III) If $f$ has Type* III*, then there are two edges $e_1$ and $e_2$. Assume the length of $e_1$ is less than or equal to that of $e_2$. The unmarked edge lengths are given by*

$$\begin{aligned} L(e_1) &= \min\left(\frac{1}{2}\left(\mathrm{val}(I_{18}) - \frac{9}{2}\,\mathrm{val}(I_4)\right), \frac{1}{4}\left(\mathrm{val}(\Delta) - 2\,\mathrm{val}(I_4)\right)\right), \\ L(e_2) &= \mathrm{val}(\Delta) - 2\,\mathrm{val}(I_4) - 2L(e_1). \end{aligned} \tag{7.1}$$

*If $(q, \ell)$ has Type* III.1*, then we write $e_1$ for the edge adjacent to the marked point and $e_2$ for the other edge. The edge lengths are then given by*

$$\begin{aligned} L(e_1) &= \frac{1}{10}(5\,\mathrm{val}(j_2) - 2\,\mathrm{val}(j_5)), \\ L(e_2) &= \frac{1}{2}(\mathrm{val}(\Delta) - 2\,\mathrm{val}(I_4)) - L(e_1). \end{aligned}$$

*For trees of Type* III.2*, they are as in (7.1).*

Combining these theorems, we then immediately obtain a description of the reduction types of Picard curves in terms of quintic and $(4, 1)$-invariants.

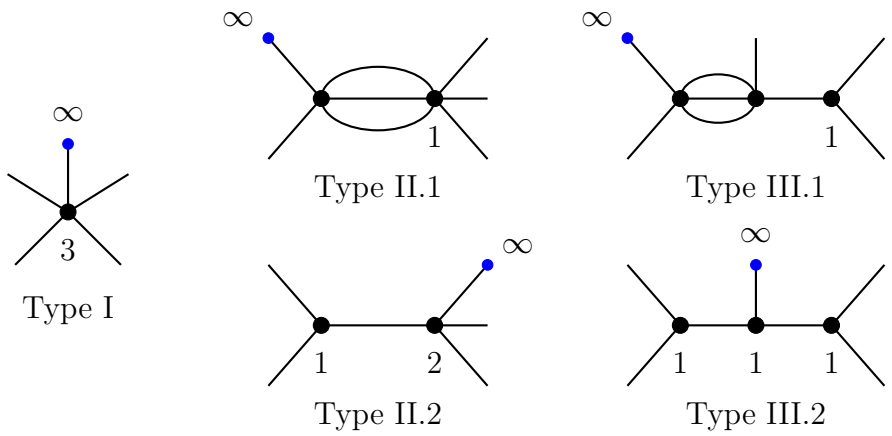



Figure 7.3: Reduction types of Picard curves

**Corollary 7.4** (**R**eduction types of Picard curves)**.** *The reduction type of the Picard curve* $y^3\ell(x,z) = q(x,z)$ *is completely determined by the tropical invariants of the* $(4,1)$*-form* $(q,\ell)$*. The tree types given in Figure 7.2 correspond to the reduction types in Figure 7.3.*

We thus obtain a description of the moduli space of tropical Picard curves in terms of invariants of quintics and $(4,1)$-forms. The work in [31] shows that these invariants in fact give rise to Picard modular forms. This is completely analogous to the cases $n = 4$ and $n = 6$ mentioned above. For instance, for $n = 4$ we have the invariants $c_4$ and $c_6$ corresponding to Eisenstein modular forms for $\mathrm{SL}(2,\mathbb{Z})$. The $j$-invariant is a rational function in terms of these modular forms and the classical criterion an elliptic curve $E$ has (potential) good reduction if and only if its $j$-invariant is non-negative expresses the tropical moduli space in terms of these modular forms. For $n = 6$ the Igusa invariants similarly give rise to modular forms (see [78, Section 6]) and the criteria given in [115] and [89] again express the tropical moduli space in terms of modular forms. We view our results as extensions of those for Picard curves.

The results obtained in this chapter can be seen as a natural continuation of [115] and [89]. In the first, criteria for the seven reduction types of curves of genus two were given in terms of the Igusa invariants. In [89], this result was extended to arbitrary complete non-archimedean fields and an easier proof was given. This chapter in turn was based on [87], where skeleta of general superelliptic curves are studied. In the latter, it was shown that one can recover the skeleton from tropicalizations of certain functions in the coefficients of $f(x)$. The key difference between this chapter and the latter is that the functions we give here are projective invariants of binary forms. This can be used to interpret the criteria in terms of Picard modular forms, giving a stronger connection to various moduli spaces in the literature.

We also define a general notion of a set of tropical invariants using tools from non-archimedean geometry. A notion that seems distantly related to this one appears in [99],

where spherical varieties and invariant valuations are studied. We do not restrict ourselves to spherical varieties, as we can phrase everything in terms of $G$-invariant subsets of the analytification of an algebraic variety for some group $G$. This removes any reliance on auxiliary objects such as Gröbner bases or graded algebras.

From a geometric point of view, our chapter fits into the literature as follows. Suppose that we have a group action on a variety that admits a geometric quotient. This quotient admits many possible compactifications and for every compactification we obtain a natural definition of a tropical invariant. For instance, for separable binary forms we can compactify the space using either stable binary forms (see [127]) or the symmetrized Deligne–Mumford compactification $\overline{M}_{0,n}/S_n$. We are mostly interested in the latter, since this has direct applications to reduction types of Picard curves. That is, if we write $\overline{\mathcal{N}}$ for the space of admissible $\mathbb{Z}/3\mathbb{Z}$-coverings with ramification signature $(4,1)$, then there is a natural map $\overline{\mathcal{N}} \to \overline{M}_{0,5}/S_4 \times S_1$ sending a covering to its branch locus and this map respects the boundary loci. Recent work by Cléry and van der Geer [31] shows that this map can be used to connect $(4,1)$-invariants to Picard modular forms, mirroring the classical case of elliptic curves. For future work, it would be interesting to see how tropical invariants of binary forms are connected to other moduli spaces.

Several results in this chapter were found or proved by symbolic computations. The codes and computations (implemented in SageMath [142]) are made available at

https://mathrepo.mis.mpg.de/TropicalInvariantsPicardCurves/index.html. (7.2)

This chapter is organized as follows. In Section 7.2, we review some background on invariant theory for binary forms and Picard curves. Section 7.3 introduces and discusses the notion of tropical invariants. Finally, we prove our main results, and in particular discuss the edge lengths (or thickness of singular points), in Section 7.4.

## 7.2 Background

In this section, we review the necessary background and preliminaries on invariant theory for binary forms and $(4,1)$-forms. We also recall some notions on Picard curves.

We start by recalling some facts and results from invariant theory of binary forms. For detailed treatments of invariant theory, we refer the reader to [154, 37, 135]. Fix an algebraically closed field $K$ such that $\operatorname{char}(K) = 0$.

Let $n \geq 0$ be a fixed integer. Let $A = K[a_0, \ldots, a_n]$ be the polynomial ring in $n+1$ variables. We view $A$ as a graded ring with the standard grading $\deg(a_i) = 1$. Let $V_n$ be the $A$-submodule of $A[x,z]$ consisting of homogeneous polynomials in $x$ and $z$ of total degree $n$. The reductive group $G := \mathrm{SL}(2,K)$ acts (as a right action) on the $A$-module $V_n$ as follows:

$$g^\sigma(x,z) := g(ax+bz, cx+dz), \qquad \text{for } g \in V_n \text{ and } \sigma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G.$$

**Definition 7.5.** We define the *universal* binary form $f$ of degree $n$ over $K$ as the binary form given by

$$f(x,z) = a_0x^n + a_1x^{n-1}z + \cdots + a_nz^n \in V_n.$$

A binary form over $K$ of degree $n$ is obtained by specializing the coefficients to $K$.

We obtain an action of $G$ on $A$ by sending $a_i$ to the coefficient $c_i(f^\sigma)$ of $x^{n-i}z^i$ in $f^\sigma$ for $0 \le i \le n$. This means that $G$ acts on $A$ as follows:

$$F^\sigma(a_0,\ldots,a_n) \coloneqq F(c_0(f^\sigma),\ldots,c_n(f^\sigma)), \quad \text{for } F \in A \text{ and } \sigma \in G. \tag{7.3}$$

**Example 7.6.** When $n = 2$, from the definition

$$f^\sigma(x,z) = f(ax+bz, cx+dz), \quad \text{for } \sigma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G$$

we get (see also Example 6.4):

$$\begin{bmatrix} c_0(f^\sigma) \\ c_1(f^\sigma) \\ c_2(f^\sigma) \end{bmatrix} = \begin{bmatrix} a^2 & ac & c^2 \\ 2ab & ad+bc & 2cd \\ b^2 & bd & d^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}.$$

We can then see how $\sigma$ acts on the generators $a_0, a_1, a_2$ of $A = K[a_0,a_1,a_2]$ and we have

$$a_0^\sigma = a^2a_0 + aca_1 + c^2a_2, \quad a_1^\sigma = 2aba_0 + (ad+bc)a_1 + 2cda_2 \quad \text{and } a_2^\sigma = b^2a_0 + bda_1 + d^2a_2.$$

**Definition 7.7.** A binary form is said to be *separable* if its discriminant does not vanish. A homogeneous polynomial $F \in A$ is called *$G$-invariant* if

$$F^\sigma = F, \quad \text{for all } \sigma \in G.$$

We denote the graded ring generated by all homogeneous $G$-invariant polynomials by $A^G$.

**Remark 7.8.** Notice that $\mathrm{GL}(2,K)$ acts on $A$ as in Equation (7.3), and since $K$ is algebraically closed, a homogeneous polynomial $F \in A$ is $\mathrm{SL}_2$-invariant if and only if for any $\sigma \in \mathrm{GL}_2$ we have

$$F^\sigma = \det(\sigma)^{\deg(F)}F.$$

We say that a homogeneous polynomial $F \in A$ is $\mathrm{GL}_2$-*invariant* if it satisfies the identity above. This immediately implies that the ring of invariants for $\mathrm{GL}_2$ and $\mathrm{SL}_2$ are the same. We will use these interchangeably.

Since $G = \mathrm{SL}(2,K)$ is reductive, it is a well known fact from invariant theory that the ring $A^G$ is finitely generated over $K$; see, for example, [37, Corollary 2.2.11]. To find generators of $A^G$ we can use the notion of transvectants, or *Überschiebung*, which we explain briefly

here. Let $g \in V_m$ and $h \in V_n$ with $m \geq n$ and let $r$ be an integer with $0 \leq r \leq n$. We define the bilinear map

$$\langle \cdot, \cdot \rangle_r : V_m \times V_n \to V_{m+n-2r}, \quad (g,h) \mapsto \sum_{i=0}^{r} (-1)^i \binom{r}{i} \frac{\partial^r g}{\partial^{r-i}x\, \partial^i z} \frac{\partial^r h}{\partial^i x\, \partial^{r-i} z}.$$

It turns out that this map is $G$-invariant, i.e.,

$$\langle g^\sigma, h^\sigma \rangle_r = \langle g, h \rangle_r, \qquad \text{for } (g,h) \in V_m \times V_n \text{ and } \sigma \in G.$$

The quantity $\langle g, h \rangle_r$ is called the $r$-th *transvectant* of $g$ and $h$. Notice that when $n = m = r$, the transvectant $\langle g, h \rangle_r$ has degree 0 in $x, z$ so $\langle g, h \rangle_r \in A^G$. A theorem by Gordan [81] shows that one can obtain all invariants by taking (iterated if necessary) transvectants, so this makes the set generators of $A^G$ explicit, at least in theory. In practice, the generators have only been found for binary forms of degrees up to 10; see [135, Chapter 1] or [37, Example 2.1.2].

**Example 7.9.**

1. For binary quadrics the invariant ring $A^G$ is generated by $\langle f, f \rangle_2$, which is equal to the discriminant of $f$.
2. For binary cubics, the invariant ring is again generated by the discriminant
$$\Delta = \langle \langle f, f \rangle_2, \langle f, f \rangle_2 \rangle_2.$$
3. For binary quartics $n = 4$, the invariant ring $A^G$ is generated by two algebraically independent invariants $I_2$ and $I_3$ of degrees 2 and 3 respectively. In terms of transvectants, they are given by
$$I_2 = \langle f, f \rangle_4 \quad \text{and} \quad I_3 = \langle f, \langle f, f \rangle_2 \rangle_4.$$
4. For binary quintics, the invariant ring $A^G$ is generated by four generators $I_4$, $I_8$, $I_{12}$ and $I_{18}$ of degrees 4, 8, 12 and 18, respectively. These are not algebraically independent, as $I_{18}^2$ can be expressed in terms of the other three. We refer the reader to Section 7.2 for a more detailed exposition in this case.

Now that we have introduced some basics of invariant theory, let us focus on $\mathrm{GL}_2$-invariants of binary forms. As discussed in Example 7.9, the ring of invariants for quintic binary forms is generated by four elements $I_4$, $I_8$, $I_{12}$ and $I_{18}$ of degrees 4, 8, 12 and 18, respectively. Let $f$ be the universal quintic

$$f(x,z) = a_0 x^5 + a_1 x^4 z + a_2 x^3 z^2 + a_3 x^2 z^3 + a_4 x z^4 + a_5 z^5.$$

The four invariants $I_4$, $I_8$, $I_{12}$ and $I_{18}$ can be obtained by taking transvectants of powers of $f$ in the following way:

$$\begin{aligned} I_4 &= \langle f^2, f^2 \rangle_{10}, \quad & I_8 &= \langle f^4, f^4 \rangle_{20}, \\ I_{12} &= \langle f^6, f^6 \rangle_{30}, \quad & I_{18} &= \langle \langle f^5, f^6 \rangle_{10}, f^7 \rangle_{35}. \end{aligned} \tag{7.4}$$

Note that these transvectants are different from the ones considered in [135, Section 4.4][1]. They still yield generators however, as one easily checks using the Poincaré series. The discriminant $\Delta$ of $f$ can be expressed in terms of the previous invariants as

$$\Delta = c_0 I_4^2 + c_1 I_8 \tag{7.5}$$

where $c_0, c_1$ are the constants

$$c_0 = -1/2746158938062848000000, \quad c_1 = 1/4698747464785208927059968000000.$$

Note that $\deg(\Delta) = 8$.

In Section 7.3, we will assume that $K$ is a complete non-archimedean field and associate a metric tree to a binary form over $K$. This tree is invariant under the action of $\mathrm{GL}_2$. For quintics, there are three different types, see Figure 7.1. The invariants $I_4, I_8, I_{12}, I_{18}$ and the discriminant $\Delta$ are however not enough to distinguish between the different types, as we will see in Example 7.27. In order to distinguish tree types, we need to introduce another invariant, which we call the *H-invariant*.

**Definition 7.10.** The *H-invariant* of the quintic $f = a_0x^5 + a_1x^4z + \cdots + a_5z^5$ is defined as follows:

$$H = \beta I_{12} - 396\alpha^3 I_4^3, \tag{7.6}$$

where $\alpha$ and $\beta$ are given by

$$\alpha = 2^{-17} \cdot 3^{-7} \cdot 5^{-3} \cdot 7^{-1},$$
$$\beta = 2^{-50} \cdot 3^{-27} \cdot 5^{-14} \cdot 7^{-7} \cdot 11^{-4} \cdot 13^{-3} \cdot 17^{-1} \cdot 19^{-1} \cdot 23^{-1} \cdot 29^{-1}.$$

This invariant is of degree 12 and constructed to satisfy the following property:

$$H(0, 0, 1, 0, a_4, a_5) \quad \text{is equal to the discriminant of the cubic} \quad x^3 + a_4x + a_5.$$

The motivation behind creating this invariant is as follows. For binary quintics, there are three unmarked tree types, see Figure 7.1. For the two non-trivial ones, we choose a non-trivial vertex of valency two. By applying a projective linear transformation, we can ensure that the roots are grouped as $\{\infty, \lambda_1\}$ and $\{0, 1, \lambda_2\}$. The leading coefficients of the corresponding binary form have positive valuation, so it will reduce to a cubic. The discriminant of this cubic then distinguishes between the two remaining types.

Now we shift gears to discuss invariants of $(4, 1)$-forms. Let $B = K[b_0, \ldots, b_4, c_0, c_1]$ be a polynomial ring. For an integer $n \geq 0$, we denote by $W_n$ the space of homogeneous polynomials in $B[x, z]$ of degree $n$. Here, we are interested in $(4, 1)$-binary forms, which are the elements $(q, \ell)$ of $W_4 \oplus W_1$. There is a natural group action of $G = \mathrm{SL}(2, K)$ on $(4, 1)$-forms as follows

$$(q(x, z), \ell(x, z))^\sigma = (q^\sigma(x, z), \ell^\sigma(x, z)) = (q(\sigma(x, z)), \ell(\sigma(x, z))).$$

[1]The universal binary form used in the literature is of the shape $f = \sum_{j=0}^{n} \binom{n}{j} a_j x^{n-j} z^j$. In this chapter we use $f = \sum_{j=0}^{n} a_j x^{n-j} z^j$, since in our field $K$ some binomial coefficients $\binom{n}{j}$ might have a positive valuation. Since $\mathrm{char}(K) = 0$, the invariant theory remains unchanged.

**Definition 7.11.** We define the *universal* $(4,1)$-form $(q,\ell)$ as

$$q(x,z) = b_0x^4 + b_1x^3z + \cdots + b_4z^4 \quad \text{and} \quad \ell(x,z) = c_0x + c_1z.$$

The action of $G$ on $(q,\ell)$ defines an action of $G$ on $B$ where $\sigma \in G$ acts on $B$ by sending $b_0,\ldots,b_4,c_0,c_1$ to the coefficients of $(q,\ell)^\sigma$. We denote by $B^G$ the ring of $G$-invariant polynomials in $B$. This ring is again finitely generated since $G$ is reductive. It has 5 generators $j_2,j_3,j_5,j_6$ and $j_9$ of respective degrees $2,3,5,6$ and 9, and they can be obtained using transvectants as follows:

$$\begin{aligned} j_2 &= \langle q,q\rangle_4, & j_3 &= \langle\langle q,q\rangle_2,q\rangle_4,\\ j_5 &= \langle q,\ell^4\rangle_4, & j_6 &= \langle\langle q,q\rangle_2,\ell^4\rangle_4,\\ j_9 &= \langle\langle q,\langle q,q\rangle_2\rangle_1,\ell^6\rangle_6. \end{aligned} \tag{7.7}$$

We refer the reader to [135, Section 5.4.1] for more details. See also Equation (7.2) for explicit formulas for all the invariants.

The invariants $I_4,I_8,I_{12},I_{18},\Delta,H$ for binary quintics and the invariants $j_2,j_3,j_5,j_6,j_9$ for $(4,1)$-forms are polynomials with rational coefficients. Scaling these invariants, we may assume that their coefficients are integer and are coprime; recall that $\mathrm{char}(K)=0$. The reason behind this is that we are working with a valued field $K$ in which integer scalars might have a positive valuation. This justifies the following assumption:

**Assumption 7.12.** We scale all the invariants so that their coefficients are coprime integers, and we do not change the notation. These scaled invariants are computed in (7.2), and are the ones used in our results.

Write $B_{4,1} = B$ for the coefficient ring corresponding to $(4,1)$-forms and $B_5$ for the coefficient ring corresponding to quintics. The universal $(4,1)$-form $(q,\ell)$ gives a canonical quintic $f = q\cdot\ell$. From this, we obtain an injective ring homomorphism

$$B_5 \to B_{4,1}.$$

This ring homomorphism is $G=\mathrm{SL}(2,K)$-equivariant, so we obtain an induced injective ring homomorphism

$$B_5^G \to B_{4,1}^G$$

of invariants. We identify $B_5^G$ with its image, so that we have an inclusion $B_5^G \subset B_{4,1}^G$. The invariants $I_4,I_8,I_{12}$ and $I_18$ for the quintic $f=q\ell$ can then be expressed in terms of the $j_i$

explicitly as follows:

$$\begin{aligned}
I_4 &= \frac{2}{3}j_2j_6 - \frac{1}{2}j_3j_5,\\
I_8 &= \frac{14}{9}j_2^2j_6^2 + \frac{22}{27}j_2^3j_5^2 + \frac{5}{27}j_3^2j_5^2 - \frac{14}{9}j_2j_3j_5j_6,\\
I_{12} &= \frac{4400}{243}j_2^3j_6^3 - \frac{11}{243}j_3^2j_6^3 - \frac{242}{9}j_2^2j_3j_5j_6^2 + \frac{2479}{81}j_2^4j_5^2j_6 + \frac{692}{81}j_2j_3^2j_5^2j_6\\
&\quad + \frac{7156}{243}j_2^3j_3j_5^3 - \frac{92}{243}j_3^3j_5^3,\\
I_{18} &= -\frac{625}{729}j_2^6j_5^3j_9 - \frac{512}{729}j_2^3j_3^2j_5^3j_9 - \frac{4}{729}j_2^3j_3j_6^3j_9 + \frac{1}{729}j_3^3j_6^3j_9 + \frac{1}{3}j_2^4j_3j_5^2j_6j_9\\
&\quad + \frac{4}{243}j_2^5j_5j_6^2j_9 - \frac{1}{243}j_2^2j_3^2j_5j_6^2j_9.
\end{aligned} \tag{7.8}$$

For $\Delta$ and $H$, we use Equations (7.5) and (7.6).

Finally, we summarize the invariants used in Theorems 7.1, 7.2 and 7.3, and their respective degrees as follows:

| **Invariant** | $I_4$ | $I_8$ | $I_{12}$ | $I_{18}$ | $\Delta$ | $H$ |
|---|---|---|---|---|---|---|
| **Degree** | 4 | 8 | 12 | 18 | 8 | 12 |

| **Invariant** | $j_2$ | $j_3$ | $j_5$ | $j_6$ | $j_9$ |
|---|---|---|---|---|---|
| **Degree** | 2 | 3 | 5 | 6 | 9 |

Table 7.1: Invariants of binary quintics (on the left) and (4,1)-forms (on the right) together with their degrees.

### 7.2.1 Picard curves

Let $K$ denote a field of characteristic 0. A Picard curve $X$ over $K$ is a smooth projective curve of genus 3 in $\mathbb{P}^2$ given by an equation of the form

$$y^3\ell(x,z) = q(x,z),$$

where $q$ and $\ell$ are homogeneous polynomials of degrees 4 and 1 respectively in $K[x,z]$. Here the smoothness of $X$ is equivalent to the separability of the quintic $q(x,z)\cdot\ell(x,z)$. For a Picard curve, the five distinct roots of this quintic are the branch points of the $\mathbb{Z}/3\mathbb{Z}$-covering of $\mathbb{P}^1$ induced by the rational map

$$\pi\colon X \dashrightarrow \mathbb{P}^1, \quad [x:y:z] \mapsto [x:z].$$

Either using the fact that $X$ is smooth or an explicit computation, we find that it gives a morphism, so that it is defined everywhere. The Galois group of $\pi$ is $G = \mathbb{Z}/3\mathbb{Z}$ and it acts as follows on $X$. Let $\zeta \in K$ be a primitive third root of unity. There is then an automorphism

$$\alpha : [x:y:z] \mapsto [x:\zeta y:z]$$

of order three and the quotient map $X \to X/\langle\alpha\rangle = \mathbb{P}^1$ is $\pi$. For each of the five points $P_i$ in the zero set of the quintic $q(x,z)\cdot\ell(x,z)$, there is a unique point $Q_i$ lying over $P_i$. The four points $Q_1,\dots,Q_4$ corresponding to the quartic $q(x,z)$ differ from the remaining point $Q_5$ in the sense that $\alpha$ acts by $\zeta$ on the tangent spaces of $Q_1,\dots,Q_4$ and as $\zeta^2$ on the tangent space of $Q_5$. This gives a second, and equivalent, definition of a Picard curve as a $\mathbb{Z}/3\mathbb{Z}$-covering of $\mathbb{P}^1$ branched over five points with a specified action of $\mathbb{Z}/3\mathbb{Z}$ on the tangent spaces of the ramification points; see [3]. This rigidification can also be given in terms of differential forms, see [31, Section 5].

## 7.3 Tropical invariants for general group actions

In this section, we define the notion of a set of tropical invariants for general group actions. The main underlying idea is that they separate orbits that have the same limit points in some toric compactification. For binary forms, this gives us the notion of a set of tropical invariants. For Picard curves, there are two interesting types of binary forms: quintics and $(4,1)$-forms. We will see in Section 7.4 that we can give a tropical set of invariants for these forms.

We assume throughout this section that $K$ is a complete, non-archimedean and algebraically closed field $K$ of characteristic 0 with valuation ring $\mathcal{O}_K$ and residue field $k$.

### 7.3.1 Compactifications and tropicalizations

Let $U$ be an irreducible variety over $K$ with Berkovich analytification $U^{\text{an}}$. We can define tropical separating sets for partitions of subsets of $U^{\text{an}}$ as follows.

**Definition 7.13.** Let $U^{\text{an}} \supset S = \bigsqcup S_i$ be a partition of a subset $S$ of $U^{\text{an}}$ into disjoint subsets and let $\phi : U \to X(\Delta)$ be an embedding of $U$ into a toric variety $X(\Delta)$ with fan $\Delta$. We say that $\phi$ is *separating* for the partition if $\operatorname{trop}(\phi(S_i)) \cap \operatorname{trop}(\phi(S_j)) = \emptyset$ for $i \neq j$, where $\operatorname{trop}(\cdot)$ is the natural tropicalization map associated to the toric variety $X(\Delta)$. If the embedding is given by the global sections of a line bundle on $U$ and $\phi$ is separating for a partition, then we call the sections a *tropical separating set* for the partition. If the sections of the line bundle give a morphism to $X(\Delta) = \mathbb{P}(a_1,\dots,a_n)$, then we say that they are a *projective tropical separating set*.

**Example 7.14.** If we take $S = \{P_1\} \cup \{P_2\}$ for two points $P_1, P_2 \in U^{\text{an}}$, then the proof of [131, Theorem 1.1] shows that we can find an embedding that separates these points.

The partitions of $U^{\text{an}}$ we are interested in are given using compactifications of $U$ as follows. Let $\mathcal{X} \to \operatorname{Spec}(R)$ be a proper flat $\mathcal{O}_K$-scheme and let $U \to \mathcal{X}$ be an open immersion. Consider the reduction map

$$\operatorname{red} : U^{\text{an}} \to \mathcal{X}_s \tag{7.9}$$

as in [84, Section 4], where $\mathcal{X}_s$ is the special fiber of $\mathcal{X}$. We now obtain a partition of $U^{\mathrm{an}}$ by taking the inverse image under $\mathrm{red}(\cdot)$ of a partition of $\mathcal{X}_s$.

### 7.3.2 Tropical invariants

We start with the definition of tropical weighted projective space.

**Definition 7.15.** Let $(\mathbb{TA}^n)^* = \overline{\mathbb{R}}^n \setminus \{(\infty, \dots, \infty)\}$, the punctured tropical affine space. For a fixed $n$-tuple of natural numbers $(a_1, \dots, a_n)$, we define an action of $\mathbb{R}$ on $(\mathbb{TA}^n)^*$ as follows:

$$\lambda \odot (x_1, \dots, x_n) \coloneqq (x_1, \dots, x_n) + \lambda \cdot (a_1, \dots, a_n).$$

The (set-theoretic) quotient of $(\mathbb{TA}^n)^*$ by this action is called the *tropical weighted projective space*, and is denoted by $\mathbb{TP}(a_1, \dots, a_n)$.

Now, let $U$ be an irreducible variety over $K$ and let $G$ be a group scheme over $K$ acting on $U$. Suppose that there exists a geometric quotient $V = U/G$ for $U$ and $G$. The notion of a tropical separating from Section 7.3.1 now allows us to define a set of tropical invariants.

**Definition 7.16.** Let $V^{\mathrm{an}} \supset S = \bigsqcup S_i$ be a partition of a subset $S$ of $V^{\mathrm{an}}$. A set of *tropical invariants* for the partition is a tropical separating set for the partition. If the tropical invariants give an embedding into a weighted projective space $\mathbb{P}(a_1, \dots, a_n)$, then we call this a set of *tropical projective invariants*. The image of the tropicalization is the space $\mathbb{TP}(a_1, \dots, a_n)$.

Since points of $V^{\mathrm{an}}$ correspond to orbits in $U^{\mathrm{an}}$, we find that a set of tropical invariants separates a partition of a subset of $U^{\mathrm{an}}$ that is stable under the action of $G$.

**Example 7.17.** Consider the ring $K[x, y]$. The group symmetric $S_2 \simeq \mathbb{Z}/2\mathbb{Z}$ acts on $K[x, y]$ by $x \mapsto y$ and $y \to x$. The invariant ring is $K[xy, x + y]$ and, on the level of schemes, the map $\mathrm{Spec}(K[x, y]) \to \mathrm{Spec}(K[xy, x + y])$ gives the geometric quotient. Consider the standard embedding of $\mathbb{A}^2 = \mathrm{Spec}(K[x, y])$ into $\mathbb{P}^2$. The group action then extends to $\mathbb{P}^2$. The boundary is isomorphic to $\mathbb{P}^1$ and we decompose $\mathbb{P}^2$ into

$$\overline{S}_0 = \{[x : y : z] \mid z \neq 0\},\ \overline{S}_1 = \{[1 : 1 : 0]\} \text{ and } \overline{S}_2 = \{[x : y : 0] \mid x \neq y\}.$$

This induces a partition

$$S_0 \sqcup S_1 \sqcup S_2 = (K[xy, x + y])^{\mathrm{an}}$$

since the group action preserves the $\overline{S_i}$. We now note that the standard embedding

$$\mathrm{Spec}(K[xy, x + y]) = \mathbb{A}^2 \to \mathbb{P}^2$$

is not tropically separating since $\mathrm{char}(K) \neq 2$. If we however consider the embedding into $\mathbb{P}^3$ given by $\{xy, x + y, (x - y)^2, 1\}$, then this does form a tropical set of invariants.

We will see many more examples of tropical invariants in Sections 7.3.3 and 7.3.4.

### 7.3.3 Moduli of tropical curves

We now review moduli of stable curves of genus zero. The functor $\mathcal{M}_{0,n} : \mathbf{Sch} \to \mathbf{Sets}$ from the category of schemes to the category of sets, defined by

$$\mathcal{M}_{0,n}(S) = \{\text{smooth curves } C \to S \text{ of genus zero with } n \text{ marked points}\}/\sim \tag{7.10}$$

for $n \geq 3$ is representable by a scheme $M_{0,n}$. We refer to this as the moduli space of smooth $n$-marked curves of genus zero. There is a natural compactification of this space, given by replacing smooth curves by stable curves in (7.10). We denote the corresponding scheme by $\overline{M}_{0,n}$. It is proper and flat over $\mathrm{Spec}(\mathbb{Z})$. The boundary locus $\Delta = \overline{M}_{0,n} \backslash M_{0,n}$ corresponds to non-smooth stable curves.

The symmetric group $S_n$ acts on the above moduli spaces by permuting the marked points. Since $S_n$ is finite and $M_{0,n}$ and $\overline{M}_{0,n}$ are quasi-projective, there exist geometric quotients $M_{0,n}/S_n$ and $\overline{M}_{0,n}/S_n$ which fit into a commutative diagram

$$\begin{array}{ccc} M_{0,n} & \longrightarrow & \overline{M}_{0,n} \\ \downarrow & & \downarrow \\ M_{0,n}/S_n & \longrightarrow & \overline{M}_{0,n}/S_n \end{array}$$

The scheme $\overline{M}_{0,n}/S_n$ is again proper over $\mathbb{Z}$. The geometric points of $M_{0,n}/S_n$ (resp. $\overline{M}_{0,n}/S_n$) can be identified with smooth (resp. stable) curves of genus 0 with $n$ unordered points.

We can also describe their abstract tropicalizations as in [2, Section 4]. We start with the original set-up without quotients. Consider phylogenetic trees with $n$ marked leaves, together with a length function $\ell$ on the non-leaves. These are the points of the tropical moduli space $M_{0,n}^{\mathrm{trop}}$. It can be given the structure of a generalized cone complex as in [2, Section 2]. It can also be given as the tropicalization of the Plücker map as in [116, Theorem 6.4.12]. We obtain an abstract tropicalization map

$$\mathrm{trop} : M_{0,n}^{\mathrm{an}} \to M_{0,n}^{\mathrm{trop}}$$

by the following procedure. A point $P \in M_{0,n}^{\mathrm{an}}$ can be represented by an $L$-valued point of $M_{0,n}$, and we use the natural map

$$M_{0,n}(L) \to \overline{M}_{0,n}(L) = \overline{M}_{0,n}(R_L) \tag{7.11}$$

to obtain a stable model $\mathcal{X} \to \mathrm{Spec}(\mathcal{O}_L)$ with special fiber $\mathcal{X}_s \to \mathrm{Spec}(k_L)$. Here, $L$ is a complete valued field extension of $K$ with valuation ring $\mathcal{O}_L$ and residue field $k_L$. Let $T$ be the marked dual intersection graph of $\mathcal{X}_s$, together with the natural edge length function $\ell$ induced from $\mathcal{X}$. Note that there is a unique leaf for every marked point. We set $\mathrm{trop}(P) = (T, \ell)$.

We now define an action of $S_m$ on $M_{0,n}^{\text{trop}}$, which gives rise to a natural quotient space $M_{0,n}^{\text{trop}}/S_m$. We view an element of $M_{0,n}^{\text{trop}}$ as a metric tree $(T, \ell)$, together with an injection $i : \{1, \ldots, n\} \to L(T)$, where $L(T)$ is the set of (infinite) leaves of $T$. Let $C$ be a set of order $m \leq n$ and let $C \to \{1, \ldots, n\}$ be an injection. By permuting the leaves indexed by $C$, we then obtain an action of $S_m$ on $M_{0,n}^{\text{trop}}$, where the latter is viewed as an object in the category of generalized cone complexes. We now note that categorical quotients in the category of generalized cone complex exist since arbitrary finite colimits exist, see [2, Remark 2.6.1]. From this, we obtain the following definition of $M_{0,n}^{\text{trop}}/S_m$.

**Definition 7.18.** Let $C \to \{1, \ldots, n\}$ be a given injection inducing an action of $S_m$ on $M_{0,n}^{\text{trop}}$. The space $M_{0,n}^{\text{trop}}/S_m$ is the categorical quotient of $M_{0,n}^{\text{trop}}$ under the action of $S_m$.

Suppose that $C = \{1, \ldots, n\}$. Set-theoretically, the points of $M_{0,n}^{\text{trop}}/S_n$ correspond to unmarked phylogenetic metric trees with $n$ leaves. As a topological space, it has a natural stratification in terms of unmarked tree types. A similar interpretation also holds for other markings.

**Example 7.19.** There are three non-trivial marked tree types for $n = 4$, giving three cones $\mathbb{R}_{\geq 0}$. The types with edge lengths zero are all identified, so we glue these cones together to obtain $M_{0,4}^{\text{trop}}$, which is a standard tropical line. The three cones are permuted by $S_4$, giving the quotient $M_{0,4}^{\text{trop}}/S_4 = \mathbb{R}_{\geq 0}$.

**Example 7.20.** Let $n = 5$. Then, there are exactly three unmarked types: I, II and III; they are depicted in Figure 7.1. Type III corresponds to a folded positive orthant of dimension 2, since the automorphism group of the underlying graph is $\mathbb{Z}/2\mathbb{Z}$. Figure 7.4 represents the space $M_{0,5}^{\text{trop}}/S_5$.

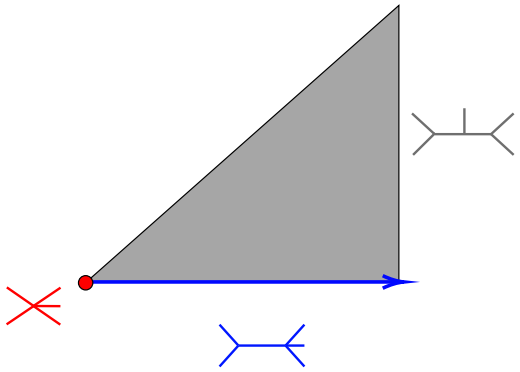


Figure 7.4: The space $M_{0,5}^{\text{trop}}/S_5$

**Example 7.21.** Consider the natural injection $C = \{1, \ldots, 4\} \to \{1, \ldots, 5\}$, giving rise to an action of $S_4$ on $M_{0,5}^{\text{trop}}$. We view the quotient $M_{0,5}^{\text{trop}}/S_4$, which is represented in Figure 7.5, as the moduli space of phylogenetic trees with 4 unmarked leaves and 1 marked leaf. There

are five different corresponding phylogenetic tree types: I, II.1, II.2, III.1 and III.2; see Figure 7.2.

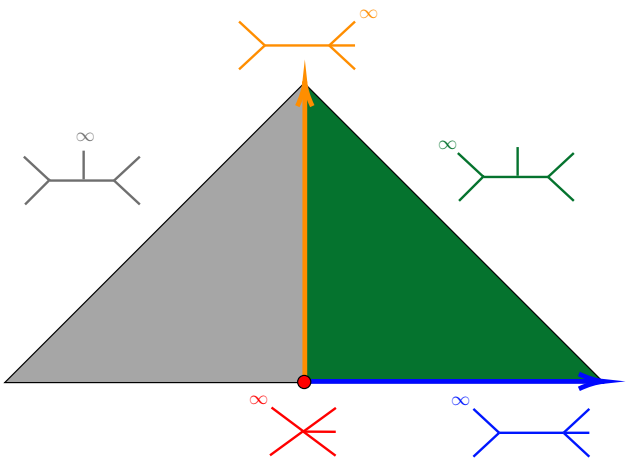



Figure 7.5: The space $M_{0,5}^{\mathrm{trop}}/S_4$

In Figures 7.4 and 7.5, the (possibly folded) positive orthants are glued with respect to degeneration of the corresponding tree types, which is shown in Figure 7.6. We refer to [119, Section 2] and [30, Section 4] for more details.

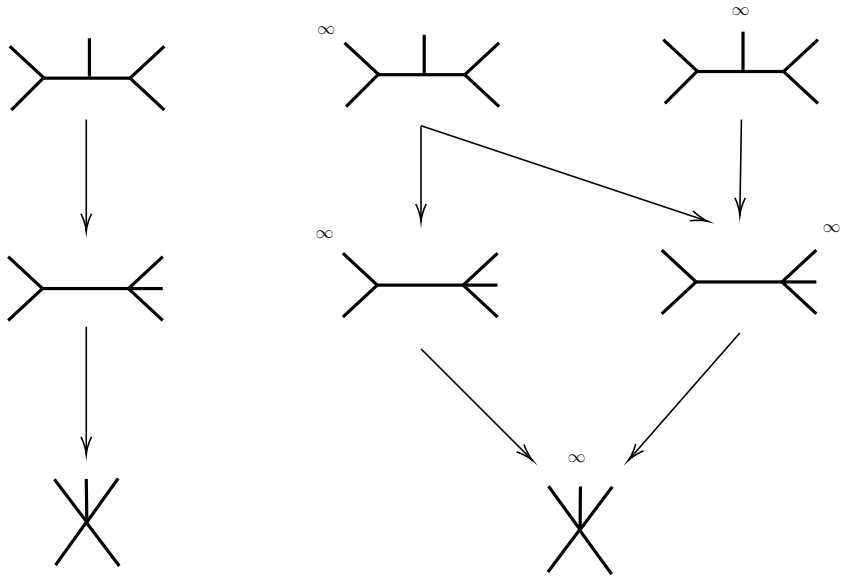



Figure 7.6: Degenerations of trees

We can again introduce an abstract tropicalization map as follows. An $L$-valued point of $M_{0,n}^{\mathrm{an}}/S_n$ for $L$ a valued field corresponds to a smooth curve of genus 0 with $n$ unordered points. We then obtain a phylogenetic metric tree as before, but we now forget the markings, giving an element of $M_{0,n}^{\mathrm{trop}}/S_n$. All in all, this gives a map

$$\mathrm{trop} : M_{0,n}^{\mathrm{an}}/S_n \to M_{0,n}^{\mathrm{trop}}/S_n. \tag{7.12}$$

**Remark 7.22.** The map in Equation (7.12) in particular gives a partition of $M_{0,n}^{\mathrm{an}}/S_n$ by considering the inverse images of the loci of unmarked tree types. We will use this partition to obtain a partition of the space of binary forms in Section 7.3.4. We note that the partition of $M_{0,n}^{\mathrm{an}}/S_n$ is a special instance of the partitions we considered in Section 7.3.1 since the construction in Equation (7.11) gives the reduction map from Equation (7.9).

### 7.3.4 Trees and tropical binary forms

Let $V = V_{k_1} \oplus \cdots \oplus V_{k_r}$ be the standard $\mathrm{SL}_2$-module and write $A^G$ for the ring of invariants. We consider $\mathrm{Proj}(A^G)$, where the grading is induced by the degree of an invariant. We are interested in the affine open $U = D_+(\Delta)$ for $\Delta$ the discriminant. The set of $L$-valued points $U(L)$ of $U$ can be identified with tuples of binary forms $(f_1, \ldots, f_r)$ over $L$ without common zeros and up to $\mathrm{GL}_2$-equivalence. Here $\deg(f_i) = k_i$.

We now consider an $L$-valued point of $U$. This gives a set of binary forms $(f_1, \ldots, f_r)$ over $L$. Consider the zero sets $Z(f_i)$ of the $f_i$. Note that every zero set gives a well-defined $L$-valued point of the moduli space $M_{0,k_i}/S_{k_i}$ of $k_i$ unordered points on the projective line. Indeed, taking a different equivalent binary form changes the zero set by the action of $\mathrm{GL}_2$, so that the induced map

$$(\mathbb{P}^1, Z(f)) \to (\mathbb{P}^1, Z(f^\sigma))$$

is an isomorphism, which means that we obtain the same point in $M_{0,k_i}/S_{k_i}$. In other words, we have a set-theoretic map

$$U \to \prod M_{0,k_i}/S_{k_i}.$$

The partition on the Berkovich analytification of $\prod M_{0,k_i}/S_{k_i}$ given in Remark 7.22 induces a partition of $U^{\mathrm{an}}$. We use this as our definition of a tropical invariant of a binary form.

**Definition 7.23.** A set of *tropical invariants* of a set of binary forms of degree $n$ is a tropical separating set for the partition of $U$ induced from the maps $U \to M_{0,n}^{\mathrm{an}}/S_n \to M_{0,n}^{\mathrm{trop}}/S_n$, see Equation (7.12).

For a given set of invariants $h_i$, we obtain a rational map $U \to \mathbb{P}(a_0, \ldots, a_n)$, where $\deg(h_i) = a_i$. If the $h_i$ include a set of generators of the ring of invariants, then this is automatically a morphism since the set of generators generate the nullcone. If the set of generators is a projective tropical separating set, then we call this a set of *projective tropical invariants*.

**Example 7.24.** For residue characteristics not equal to 2, the tropical invariants of a binary quartic are given by $c_4$, $c_6$ and $\Delta$; see [152, Chapter VII]. These are furthermore weighted projective invariants, giving a tropicalization map to $\mathbb{TP}(4, 6, 12)$.

**Example 7.25.** For residue characteristics not equal to 2, the tropical invariants of a binary sextic are given by the Igusa invariants from [89].

**Example 7.26.** In Section 7.4, we will show that the invariants introduced in the introduction form a set of tropical invariants for binary quintics and $(4, 1)$-forms. These then also give the reduction types of Picard curves when the residue characteristic is not 3.

As promised in Section 7.2, we now show that the invariants $I_4$, $I_8$, $I_{12}$, $I_{18}$ and $\Delta$ are not enough to distinguish between the unmarked trees of a quintic.

**Example 7.27.** Let $K = \mathbb{C}\{\{t\}\}$ be the field of Puiseux series over the complex numbers with $\mathrm{val}(t) = 1$. Consider the two quintics given by

$$f_2 = xz(x - z)(x - t^2 z)(x - 2z),$$
$$f_3 = xz(x - z)(x - tz)(x - (1 + 2t)z).$$

The tree of $f_2$ is of Type II and the tree of $f_3$ is of Type III. However, in both cases, the tropical invariants are the same:

$$[\mathrm{val}(I_4) : \mathrm{val}(I_8) : \mathrm{val}(I_{12}) : \mathrm{val}(I_{18}) : \mathrm{val}(\Delta)] = [0 : 0 : 0 : 2 : 4].$$

We thus see that we cannot distinguish between the two tree types using these invariants. This also implies that we cannot distinguish the various reduction types of Picard curves using these invariants. They are however enough to decide whether a tree is of Type I or not.

The results in this chapter show that there exists a set of projective tropical invariants for quintics and $(4, 1)$-forms. In general, we conjecture that there exists a finite set of projective tropical invariants for any set of binary forms. Moreover, there should be a practical algorithm that can calculate these tropical invariants.

## 7.4 Proofs of the main results

In this final section, we prove the main results stated in the introduction, namely Theorem 7.1, Theorem 7.2, Theorem 7.3 and Corollary 7.4. Some parts of the proofs rely on computing the invariants explicitly. The computations are made available in (7.2).

Recall that our base field $K$ is a non-archimedean, complete and algebraically closed valued field of characteristic 0 with associated data $(v, R, \mathfrak{m}, k)$. Moreover, the residue characteristic $p$ is different from $2, 3$ and $11$.

**Remark 7.28.** If the residue characteristic of $K$ is 11, then the second condition in Theorem 7.1 does not characterize trees of Type II. This is because, when $p = 11$ divides the reduction of the invariant $H$ reduces to 0 modulo $\{x \in K\colon \operatorname{val}(x) > \operatorname{val}(H)\}$, see Equation (7.13). In this case, to obtain condition for Type II, we simply need to check that the conditions for Type I and Type III are not satisfied.

### 7.4.1 Universal families

We first explain the general idea of the proofs. For a given binary quintic or $(4,1)$-form, we can apply a projective transformation $\sigma \in \mathrm{GL}(2,K)$ to rearrange the roots. This has the following effect on the tree and the invariants:

1. If $\Sigma$ is the (marked or unmarked) tree, then $\sigma$ induces an isometry $\sigma : \Sigma \to \sigma(\Sigma)$ outside the type-1 points by the results in [14, Section 2.3]. Hence the tree is unchanged.

2. The transformation $\sigma$ changes an invariant $I$ by $\det(\sigma)^{\deg(I)}$. In particular, we find that the projective tropical invariants are unchanged; see Remark 7.8.

For a given binary quintic or $(4,1)$-form, we can apply a projective transformation so that the resulting binary form is

$$f(x,z) = xz(x-z)(x-\lambda_1 z)(x-\lambda_2 z).$$

Here, we send the marked point to $\infty$. By examining the type of the tree, we then obtain various conditions on the $\lambda_i$ which we use to write down universal families for a given tree type. Here by universal, we mean that every binary form with a given tree type occurs in the given family. We then perform calculations on the universal families, which directly give the proofs for our main theorems. We can make this more precise as follows. For a given tree type, we fix a *universal algebra* $A = \mathcal{O}_K[t_i, \mu_i]$ with $\lambda_i \in A$. This contains all the necessary parameters for the universal family. We then consider $\mathcal{O}_K$-valued points $\psi : A \to \mathcal{O}_K$ such that $\operatorname{val}(\psi(t_i)) > 0$. Such an $\mathcal{O}_K$-valued point corresponds to a specific binary form of the given type. We will also write $\operatorname{val}(t_i) > 0$ if $\psi$ is understood. To calculate with reductions, we work in $A/I$, where $I = \mathfrak{m} + (t_1)$. We will also just refer to this as working modulo $\mathfrak{m}$. We note that all the invariants are elements of the universal algebra $A$. In particular, for any specialization $\psi : A \to \mathcal{O}_K$, we have that $\psi(I) \in \mathcal{O}_K$. We write $I \in A^\times$ if for every specialization $\psi$ as above, we have that $\operatorname{val}(\psi(I)) = 0$.

In the proofs of the main theorems, we will use the universal families from Table 7.2; see Figure 7.7 for the corresponding trees.

| Type | $\boldsymbol{\lambda_1}$ | $\boldsymbol{\lambda_2}$ | Conditions |
|---|---|---|---|
| I | $\mu_1$ | $\mu_2$ | $\overline{\mu}_i \neq 0, 1$ and $\overline{\mu}_1 \neq \overline{\mu}_2$ |
| II.1 | $t_1\mu_1$ | $t_1\mu_2$ | $\mathrm{val}(t_1) > 0$, $\overline{\mu}_i \neq 0$ and $\overline{\mu}_1 \neq \overline{\mu}_2$ |
| II.2 | $t_1\mu_1$ | $\mu_2$ | $\mathrm{val}(t_1) > 0$, $\overline{\mu}_i \neq 0$ and $\overline{\mu}_2 \neq 1$ |
| III.1 | $t_1\mu_1$ | $t_1t_2\mu_2$ | $\mathrm{val}(t_i) > 0$ and $\overline{\mu}_i \neq 0$ |
| III.2 | $t_1\mu_1$ | $1 + t_2\mu_2$ | $\mathrm{val}(t_i) > 0$, $\overline{\mu}_i \neq 0$ and $\mathrm{val}(t_1) \leq \mathrm{val}(t_2)$ |

Table 7.2: The chosen universal families

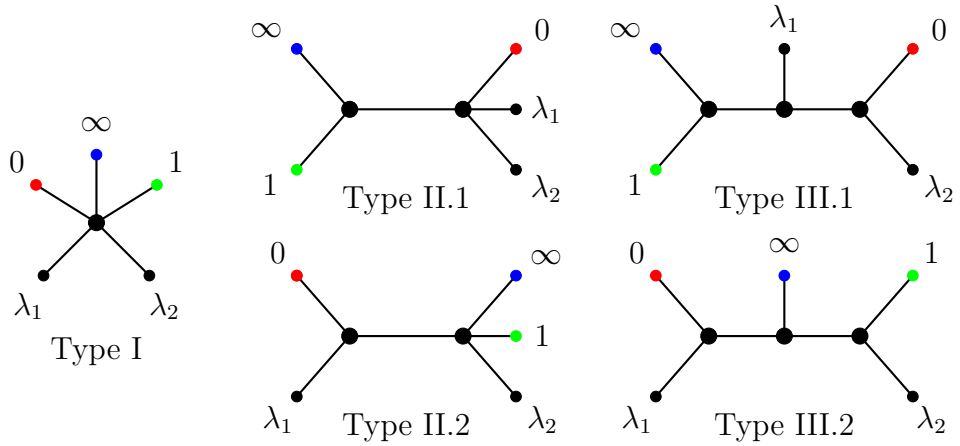



Figure 7.7: Trees corresponding the universal families in Table 7.2

### 7.4.2 Proof of Theorem 7.1

In this subsection, we prove Theorem 7.1. In the proofs, we are free to choose the universal family of either II.1 or II.2, and similarly for III.1 and III.2.

**Type I:** Computing the reduction modulo $\mathfrak{m}$ of the discriminant $\Delta$, we get

$$\overline{\Delta} = \overline{\lambda}_1^2\,\overline{\lambda}_2^2\,(\overline{\lambda}_1 - 1)^2\,(\overline{\lambda}_2 - 1)^2\,(\overline{\lambda}_1 - \overline{\lambda}_2)^2.$$

So we deduce that $\mathrm{val}(\Delta) = 0$, hence the condition $8\,\mathrm{val}(I) - \deg(I)\,\mathrm{val}(\Delta) \geq 0$ is satisfied for any $I \in S$.

**Type II:** Consider the universal family for Type II.1 in Table 7.2. We calculate the invariants $I \in S$ and find that they are divisible by $t_1^{\deg I/2}$. The reduction of $H/t_1^{\deg(H)/2}$ modulo $\mathfrak{m}$ is

$$\overline{H/t_1^{\deg(H)/2}} = -22\overline{\mu}_1^2\overline{\mu}_2^2(\overline{\mu}_1 - \overline{\mu}_2)^2. \tag{7.13}$$

Since $p \neq 2, 11$, we find that the latter is non-zero. We thus obtain

$$\operatorname{val}(H) = \deg(H)\operatorname{val}(t_1)/2 = 6\operatorname{val}(t_1),$$

and

$$\operatorname{val}(I) \geq \deg(I)\operatorname{val}(t_1)/2 \quad \text{for } I \in S \setminus \{H\}.$$

Therefore,

$$12\operatorname{val}(I) - \deg(I)\operatorname{val}(H) \geq 0 \text{ for all } I \in S.$$

On the other hand, computing the reduction modulo $\mathfrak{m}_K$ of $I_4/t_1^2$ and $I_{18}/t_1^9$, we find that

$$\begin{aligned}\overline{I_4/t_1^2} &= -2(\overline{\mu}_1^2 - \overline{\mu}_1\overline{\mu}_2 + \overline{\mu}_2^2),\\ \overline{I_{18}/t_1^9} &= -\overline{\mu}_1^2\overline{\mu}_2^2(\overline{\mu}_1 - 2\overline{\mu}_2)(\overline{\mu}_1 + \overline{\mu}_2)(2\overline{\mu}_1 - \overline{\mu}_2)(\overline{\mu}_1 - \overline{\mu}_2)^2.\end{aligned}$$

It is not so hard to check that, since $p \neq 3$, the quantities $\overline{I_4/t_1^2}$ and $\overline{I_{18}/t_1^9}$ cannot be simultaneously zero. We thus find that

$$\operatorname{val}(I_4) = 2\operatorname{val}(t_1) \quad \text{or} \quad \operatorname{val}(I_{18}) = 9\operatorname{val}(t_1).$$

Our computations give $\operatorname{val}(\Delta/t_1^6) \geq 0$, so we deduce that $\operatorname{val}(\Delta) > 4\operatorname{val}(t_1)$ and thus

$$\operatorname{val}(\Delta) - 2\operatorname{val}(I_4) > 0 \quad \text{or} \quad 9\operatorname{val}(\Delta) - 4\operatorname{val}(I_{18}) > 0.$$

Combining these, we obtain the statement of the theorem.

**Type III:** Consider the universal family for Type III.2 in Table 7.2. We compute the reduction of $I_4$ modulo $\mathfrak{m}$ to obtain

$$\overline{I}_4 = -2.$$

So, since $p \neq 2$, we deduce that $\operatorname{val}(I_4) = 0$. Computing $\Delta$ and $H$, we obtain that $\Delta$ is divisible by $t_1^2t_2^2$ and $\overline{H} = 0$. So we deduce that

$$\operatorname{val}(\Delta) - 2\operatorname{val}(I_4) > 0 \quad \text{and} \quad \operatorname{val}(H) - 3\operatorname{val}(I_4) > 0.$$

To finish the proof, we check that these conditions are sufficient. Suppose that the condition in (I) is satisfied. In particular,

$$8\operatorname{val}(I_4) - 4\operatorname{val}(\Delta) \geq 0 \text{ and } 8\operatorname{val}(I_{18}) - 18\operatorname{val}(\Delta) \geq 0.$$

But these imply that

$$\operatorname{val}(\Delta) - 2\operatorname{val}(I_4) \leq 0 \text{ and } 9\operatorname{val}(\Delta) - 4\operatorname{val}(I_{18}) \leq 0,$$

so the conditions in (II) and (III) can not be satisfied. Now suppose that the conditions in (II) are satisfied. The condition

$$\operatorname{val}(\Delta) - 2\operatorname{val}(I_4) > 0 \text{ or } 9\operatorname{val}(\Delta) - 4\operatorname{val}(I_{18}) > 0$$

gives

$$8\operatorname{val}(I_4) - \deg(I_4)\operatorname{val}(\Delta) < 0 \text{ or } 8\operatorname{val}(I_{18}) - \deg(I_{18})\operatorname{val}(\Delta) < 0,$$

and therefore, the condition in (I) can not be satisfied. On the other hand, the inequality $12\operatorname{val}(I_4) - 4\operatorname{val}(H) \geq 0$ implies $\operatorname{val}(H) - 3\operatorname{val}(I_4) \leq 0$, which means that the conditions in (III) can not be satisfied either. Finally, suppose that the conditions in (III) are satisfied. Hence,

$$8\operatorname{val}(I_4) - \deg(I_4)\operatorname{val}(\Delta) < 0 \text{ and } 12\operatorname{val}(I_4) - \deg(I_4)\operatorname{val}(H) < 0,$$

so that the conditions in (I) and (II) can not be satisfied. This finishes the proof.

### 7.4.3 Proof of Theorem 7.2

In this subsection, we prove Theorem 7.2. The strategy is the same as in Section 7.4.2: we first calculate the invariants for each universal family and show that the given inequalities hold. This gives the necessity of the conditions. The sufficiency in this case is trivial, so this concludes the proof. We start with trees of Type II.1 since the marking does not matter for trees of Type I.

**Type II.1:** We calculate $j_2$ and find that it is divisible by $t_1^2$. Computing the reductions modulo $\mathfrak{m}$ of $j_5$ we obtain $\overline{j}_5 = 1$. We then deduce that $5\operatorname{val}(j_2) - 2\operatorname{val}(j_5) > 0$.

**Type II.2:** We calculate the reductions of $j_2$ and $j_5$ modulo $\mathfrak{m}$ and find $\overline{j}_2 = \overline{\mu}_2^2$ and $\overline{j}_5 = 1$. So we deduce that $5\operatorname{val}(j_2) - 2\operatorname{val}(j_5) = 0$.

**Type III.1:** We find that $j_2$ is divisible by $t_1^2$. Calculating the reduction of $j_5$ modulo $\mathfrak{m}_K$, we obtain $\overline{j}_5 = 1$. This implies $5\operatorname{val}(j_2) - 2\operatorname{val}(j_5) > 0$.

**Type III.2:** We compute the reductions of $j_2$ and $j_5$ modulo $\mathfrak{m}_K$ and find $\overline{j}_2 = \overline{j}_5 = 1$. This implies $5\operatorname{val}(j_2) - 2\operatorname{val}(j_5) = 0$.

### 7.4.4 Proof of Theorem 7.3

For trees of Type I, there is nothing to prove.

**Type II:** We assume that the quintic $f$ has tree Type II and use the universal family for Type II.1 in Table 7.2.

We compute the invariants $\Delta, I_4$ and $I_{18}$, and find that $\Delta \in t_1^6 A$, $I_4 \in t_1^2 A$ and $I_{18} \in t_1^9 A$, where $A$ is the universal algebra in Section 7.4.1. Computing the reductions, we see

$$\begin{aligned}
\overline{\Delta/t_1^6} &= \overline{\mu}_1^2\ \overline{\mu}_2^2\ (\overline{\mu}_1 - \overline{\mu}_2)^2, \\
\overline{I_4/t_1^2} &= -2\ (\overline{\mu}_1^2 - \ \overline{\mu}_1\overline{\mu}_2 + \ \overline{\mu}_2^2), \\
\overline{I_{18}/t_1^9} &= -\overline{\mu}_1^2\ \overline{\mu}_2^2\ (\overline{\mu}_1 + \overline{\mu}_2)\ (\overline{\mu}_1 - 2\overline{\mu}_2)\ (2\overline{\mu}_1 - \overline{\mu}_2)\ (\overline{\mu}_1 - \overline{\mu}_2)^2.
\end{aligned}$$

Notice that, since $p \neq 3$, the two quantities $\overline{I_4/t_1^2}$ and $\overline{I_{18}/t_1^9}$ cannot vanish simultaneously. So we get

$$\operatorname{val}(\Delta) = 6\operatorname{val}(t_1) \quad \text{and} \quad (\ \operatorname{val}(I_4) = 2\operatorname{val}(t_1) \quad \text{or} \quad \operatorname{val}(I_{18}) = 9\operatorname{val}(t_1)\ )\,.$$

So we deduce that

$$\operatorname{val}(t_1) = \frac{1}{2}\left(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)\right) \quad \text{or} \quad \operatorname{val}(t_1) = \frac{1}{3}\left(2\operatorname{val}(\Delta) - \operatorname{val}(I_{18})\right),$$

and the valuation of $t_1$, the length of the unique non-trivial edge, is then the maximum of these two quantities, i.e.,

$$L(e_1) = \max\left(\frac{1}{2}\left(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)\right),\ \ \frac{1}{3}\left(2\operatorname{val}(\Delta) - \operatorname{val}(I_{18})\right)\right).$$

**Type III:** Before we give the proof of Theorem 7.3 for trees of Type III, we shortly discuss the various formulas occurring in that theorem. For a tree of Type III, we can only recover the edge lengths from the quintic invariants up to a permutation of the edges, see Example 7.20. A set of representatives of the edges here is given by $(e_1, e_2)$ with $L(e_1) \leq L(e_2)$. For trees of Type III.2, this symmetry continues to hold, so the formulas do not change. For trees of Type III.1, there is no such symmetry, meaning that we can single out the edge next to the marked point. See Section 7.4.4 for the formulas in this case.

Now suppose that $f$ has tree Type III. Computing the invariants $\Delta, I_4$ and $I_{18}$, we obtain[2]

$$\begin{aligned}
\Delta = {} & \boldsymbol{t_1^2\ t_2^2}\ \mu_1^2\ \mu_2^2\ (t_2\mu_2 + 1)^2\ (-t_1\mu_1 + t_2\mu_2 + 1)^2\ (t_1\mu_1 - 1)^2, \\
I_{18} = {} & \boldsymbol{(\mu_1 t_1 - \mu_2 t_2)(\mu_1 t_1 + \mu_2 t_2)}(-\mu_1 t_1 + \mu_2 t_2 + 2)(\mu_2^2 t_2^2 - \mu_1 t_1 + 2\mu_2 t_2 + 1) \\
& (\mu_2^2 t_2^2 + \mu_1 t_1 - 1)(-2\mu_1\mu_2 t_1 t_2 + \mu_2^2 t_2^2 - \mu_1 t_1 + 2\mu_2 t_2 + 1)(\mu_1\mu_2 t_1 t_2 - \mu_2 t_2 - 1) \\
& (\mu_1\mu_2 t_1 t_2 + \mu_2 t_2 + 1)(\mu_1\mu_2 t_1 t_2 - \mu_1 t_1 + 1)(\mu_1\mu_2 t_1 t_2 + \mu_1 t_1 - 1) \\
& (\mu_1\mu_2 t_1 t_2 + \mu_1 t_1 - 2\mu_2 t_2 - 1)(\mu_1\mu_2 t_1 t_2 + 2\mu_1 t_1 - \mu_2 t_2 - 1)(-\mu_1^2 t_1^2 + \mu_2 t_2 + 1) \\
& (-\mu_1^2 t_1^2 + 2\mu_1\mu_2 t_1 t_2 + 2\mu_1 t_1 - \mu_2 t_2 - 1)(\mu_1^2 t_1^2 - 2\mu_1 t_1 + \mu_2 t_2 t_2 + 1)
\end{aligned}$$

[2]The bold-faced factors have positive valuation, the remaining factors have valuation 0.

and $I_4 \in A$ with $\overline{I}_4 = -2$. So we deduce that

$$\operatorname{val}(\Delta) = 2\operatorname{val}(t_1) + 2\operatorname{val}(t_2) \quad \text{and} \quad \operatorname{val}(I_4) = 0.$$

If $\operatorname{val}(t_1) < \operatorname{val}(t_2)$, then we get $\operatorname{val}(I_{18}) = 2\operatorname{val}(t_1)$ and this gives

$$L(e_1) = \operatorname{val}(t_1) = \frac{1}{2}\left(\operatorname{val}(I_{18}) - \frac{9}{2}\operatorname{val}(I_4)\right).$$

If, on the other hand, $\operatorname{val}(t_1) = \operatorname{val}(t_2)$, then we have

$$L(e_1) = \operatorname{val}(t_1) = \frac{1}{4}\left(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)\right).$$

Therefore, in both cases, we get

$$L(e_1) = \operatorname{val}(t_1) = \min\left(\frac{1}{2}\left(\operatorname{val}(I_{18}) - \frac{9}{2}\operatorname{val}(I_4)\right), \frac{1}{4}\left(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)\right)\right).$$

The length of the second edge is $\operatorname{val}(t_2)$ and can be computed using $\Delta$ as

$$L(e_2) = \operatorname{val}(t_2) = \frac{1}{2}(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)) - \operatorname{val}(t_1).$$

Hence we deduce that

$$L(e_1) = \min\left(\frac{1}{2}\left(\operatorname{val}(I_{18}) - \frac{9}{2}\operatorname{val}(I_4)\right), \frac{1}{4}\left(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)\right)\right),$$
$$L(e_2) = \frac{1}{2}(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)) - L(e_1).$$

**Type III.1:** Now let $(q, \ell)$ be a $(4,1)$-form with tree Type III.1. Let $e_1$ be the edge adjacent to $\infty$ and $e_2$ the second edge in the tree Type III.1, see Figure 7.2. Using the universal family for Type III.1 and computing the invariants we find $j_2 \in t_1^2 A^\times$, $j_5 = 1$, $\Delta \in t_1^6 t_2^2 A$ and $I_4 \in t_1^2 A^\times$ and

$$\overline{j_2/t_1^2} = \overline{\mu}_1^2, \quad \overline{\Delta/(t_1^6 t_2^2)} = \overline{\mu}_1^4 \overline{\mu}_2^2, \quad \text{and} \quad \overline{I_4/t_1^2} = -2\overline{\mu}_1^2.$$

So we deduce that

$$L(e_1) = \operatorname{val}(t_1) = \frac{1}{10}(5\operatorname{val}(j_2) - 2\operatorname{val}(j_5)),$$

and

$$L(e_2) = \operatorname{val}(t_2) = \frac{1}{2}\left(\operatorname{val}(\Delta) - 2\operatorname{val}(I_4)\right) - L(e_1).$$

### 7.4.5 Proof of Corollary 7.4

We now recall from [87] how the reduction type of a Picard curve $y^3\ell(x,z) = q(x,z)$ can be recovered from the $(4,1)$-marked tree of $(q,\ell)$. We refer the reader to [87, Section 1.2] for the definition of the reduction type of a curve. We note here that our assumption that $K$ is algebraically closed is not restrictive. Namely, if we are interested in the reduction type of a curve over a complete discretely valued field $K$, then its reduction type is completely determined by the reduction type of the base change over $\overline{K}$, see [87, Remark 3.6]. Finally, we note that the notion of an edge length used here is the same as the notion of *thickness* used in other sources.

Let $X$ be a Picard curve over $K$. The branch locus $B$ of the covering

$$X \to \mathbb{P}^1$$

given by $[x:y:z] \mapsto [x:z]$ is the zero locus of $q \cdot \ell$. The minimal skeleton of the marked curve $(\mathbb{P}^{1,\mathrm{an}}, B)$ is then the $(4,1)$-marked tree of $(q,\ell)$. By applying a projective transformation, we can assume that the zero of $\ell$ is $\infty$.

For tame coverings, we have that the inverse image of a skeleton is a skeleton (see [87, Theorem 3.1] or [88, Theorem 1.1]), so we obtain a map

$$\Sigma' \to \Sigma,$$

where $\Sigma$ is the $(4,1)$-marked tree and $\Sigma'$ is its inverse image under the morphism of Berkovich analytifications $X^{\mathrm{an}} \to \mathbb{P}^{1,\mathrm{an}}$. Consider the dehomogenized polynomial $q = q(x,1)$. The criteria in [87, Section 3.1] allow us to reconstruct $\Sigma'$ explicitly in terms of the piecewise-linear function $-\log|q|$ on $\mathbb{P}^{1,\mathrm{an}} \backslash B$. This function can in turn be obtained from potential theory. We then find that over an edge in $\Sigma$, there are three edges if and only if the slope of $-\log|q|$ is divisible by three. If it is not divisible by three, then the length of an edge has to be divided by three. That is, the expansion factor $d_{e'/e}$ in this case is three. This data is enough to determine the skeleton for Picard curves, as the weights of the vertices are determined by the Riemann–Hurwitz conditions. The resulting graphs can be found in Figure 7.3.

*Proof of Corollary 7.4.* By Theorem 7.2, the $(4,1)$-marked tree type of $(q,\ell)$ is determined by the tropical invariants. The edge lengths of the $(4,1)$-marked tree type are then given by Theorem 7.3. To obtain the edge lengths for the curve $X$, we use the formula

$$d_{e'/e} L(e') = L(e),$$

where $d_{e'/e}$ is the expansion factor, see [8, Definition 2.4, Theorem 4.23]. □

## 7.5 Conclusion

To conclude, the results discussed in this chapter give a method to determined the tree type of a binary quintic or a $(4,1)$-form from the valuation of their associated invariants.

Every tree type corresponds to a cone cut out by a set of inequalities in the valuations of the invariants. This has an application in terms of determining reduction types of a Picard curve in terms of the invariants of the $(4, 1)$-form involved in its equation.

# Bibliography


[1] Peter Abramenko and Kenneth S. Brown. *Buildings*. Vol. 248. Graduate Texts in Mathematics. Theory and applications. Springer, New York, 2008, pp. xxii+747 (cit. on pp. 10, 21, 28, 112, 113).

[2] Dan Abramovich, Lucia Caporaso, and Sam Payne. "The tropicalization of the moduli space of curves". In: *Ann. Sci. Éc. Norm. Supér. (4)* 48.4 (2015), pp. 765–809 (cit. on pp. 155, 156).

[3] Jeffrey D. Achter and Rachel Pries. "The integral monodromy of hyperelliptic and trielliptic curves". In: *Mathematische Annalen* 338.1 (Dec. 2006), pp. 187–206 (cit. on p. 153).

[4] Takashi Agoh. "On Fermat's last theorem and the Bernoulli numbers". In: *Journal of Number Theory* 15.3 (1982), pp. 414–422 (cit. on p. 69).

[5] Rida Ait El Manssour and Antonio Lerario. "Probabilistic enumerative geometry over $p$-adic numbers: linear spaces on complete intersections". In: *Annales Henri Lebesgue* (2022). To appear (cit. on pp. 129, 133).

[6] Marianne Akian, Stéphane Gaubert, and Alexander Guterman. "Tropical polyhedra are equivalent to mean payoff games". In: *Internat. J. Algebra Comput.* 22.1 (2012), pp. 1250001, 43 (cit. on p. 19).

[7] Xavier Allamigeon et al. "What tropical geometry tells us about the complexity of linear programming". In: *SIAM Rev.* 63.1 (2021), pp. 123–164 (cit. on p. 20).

[8] Omid Amini et al. "Lifting harmonic morphisms I: metrized complexes and Berkovich skeleta". In: *Res. Math. Sci.* 2 (2015), Art. 7, 67 (cit. on p. 166).

[9] Hans C. Andersen and Persi Diaconis. "Hit and run as a unifying device". In: *J. Soc. Fr. Stat. & Rev. Stat. Appl.* 148.4 (2007), pp. 5–28 (cit. on p. 42).

[10] Tsuneo Arakawa, Tomoyoshi Ibukiyama, and Masanobu Kaneko. *Bernoulli numbers and zeta functions*. Springer Monographs in Mathematics. With an appendix by Don Zagier. Springer, Tokyo, 2014, pp. xii+274 (cit. on p. 69).

[11] Benjamin Assarf et al. "Computing convex hulls and counting integer points with `polymake`". In: *Math. Program. Comput.* 9.1 (2017), pp. 1–38 (cit. on p. 107).

[12] David Avis. "On the extreme rays of the metric cone". In: *Canadian J. Math.* 32.1 (1980), pp. 126–144 (cit. on p. 106).

[13] Raymond Ayoub. "Euler and the zeta function". In: *Amer. Math. Monthly* 81 (1974), pp. 1067–1086 (cit. on p. 69).

[14] Matthew Baker, Sam Payne, and Joseph Rabinoff. "On the structure of non-Archimedean analytic curves". In: vol. 605. Contemp. Math. Amer. Math. Soc., Providence, RI, 2013, pp. 93–121 (cit. on p. 160).

[15] Elizabeth Baldwin and Paul Klemperer. "Tropical geometry to analyse demand". In: *https://www.nuff.ox.ac.uk/economics/papers/2012/tropicalgeom.pdf* (2013). Unpublished paper (cit. on p. 19).

[16] Manjul Bhargava et al. "The density of polynomials of degree $n$ over $\mathbb{Z}_p$ having exactly $r$ roots in $\mathbb{Q}_p$". In: *Proceedings of the London Mathematical Society* 124.5 (2022), pp. 713–736 (cit. on pp. 41, 129, 133).

[17] Philippe Biane, Jim Pitman, and Marc Yor. "Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions". In: *American Mathematical Society. Bulletin. New Series* 38.4 (2001), pp. 435–465 (cit. on p. 69).

[18] Albert Kh. Bikulov and Igor' Vasil'evich Volovich. "$p$-adic Brownian motion". In: *Izv. Ross. Akad. Nauk Ser. Mat.* 61.3 (1997), pp. 75–90 (cit. on p. 41).

[19] Armand Borel. *Linear algebraic groups*. Second. Vol. 126. Graduate Texts in Mathematics. Springer-Verlag, New York, 1991, pp. xii+288 (cit. on p. 63).

[20] Wieb Bosma, John Cannon, and Catherine Playoust. "The Magma algebra system. I. The user language". In: *J. Symbolic Comput.* 24.3-4 (1997). Computational algebra and number theory (London, 1993), pp. 235–265 (cit. on p. 64).

[21] Nicolas Bourbaki. *Lie groups and Lie algebras. Chapters 1–3*. Elements of Mathematics (Berlin). Translated from the French, Reprint of the 1975 edition. Springer-Verlag, Berlin, 1989, pp. xviii+450 (cit. on p. 69).

[22] Nicolas Bourbaki. *Lie groups and Lie algebras. Chapters 4–6*. Elements of Mathematics (Berlin). Translated from the 1968 French original by Andrew Pressley. Springer-Verlag, Berlin, 2002, pp. xii+300 (cit. on pp. 13, 113).

[23] Paul Breiding and Orlando Marigliano. "Random points on an algebraic manifold". In: *SIAM J. Math. Data Sci.* 2.3 (2020), pp. 683–704 (cit. on pp. 14, 41).

[24] Jan Hendrik Bruinier. "Hilbert modular forms and their applications". In: *The 1-2-3 of modular forms*. Universitext. Springer, Berlin, 2008, pp. 105–179 (cit. on p. 65).

[25] Urtzi Buijs, José G. Carrasquel-Vera, and Aniceto Murillo. "The gauge action, DG Lie algebras and identities for Bernoulli numbers". In: *Forum Math.* 29.2 (2017), pp. 277–286 (cit. on p. 69).

[26] Peter Bürgisser, Avinash Kulkarni, and Antonio Lerario. "Nonarchimedean integral geometry". In: *arXiv:2206.03708* (2022) (cit. on pp. 41, 47).

[27] Dustin Cartwright et al. "Mustafin varieties". In: *Selecta Math. (N.S.)* 17.4 (2011), pp. 757–793 (cit. on p. 13).

[28] Xavier Caruso. "Where are the zeroes of a random $p$-adic polynomial?" In: *Forum Math. Sigma* 10 (2022), Paper No. e55, 41 (cit. on pp. 19, 41, 129, 133).

[29] John William Scott Cassels. *Local fields*. Vol. 3. Cambridge University Press Cambridge, 1986 (cit. on p. 4).

[30] Melody Chan. "Lectures on tropical curves and their moduli spaces". In: *Moduli of curves*. Vol. 21. Lect. Notes Unione Mat. Ital. Springer, Cham, 2017, pp. 1–26 (cit. on p. 157).

[31] Fabien Cléry and Gerard van der Geer. "Generating Picard modular forms by means of invariant theory". In: *arXiv:2110.00849* (2021) (cit. on pp. 146, 147, 153).

[32] Fabien Cléry and Gerard van der Geer. "Modular Forms of Degree 2 and Curves of Genus 2 in Characteristic 2". In: *Int. Math. Res. Not.* 7 (2022), pp. 5204–5218 (cit. on p. 2).

[33] Alexander Clifton et al. "Continuously Increasing Subsequences of Random Multiset Permutations". In: *arXiv:2110.10315* (2021) (cit. on pp. 2, 14, 71, 89, 91, 92, 96).

[34] Carlos A. Coelho. "The wrapped gamma distribution and wrapped sums and linear combinations of independent gamma and Laplace distributions". In: *J. Stat. Theory Pract.* 1.1 (2007), pp. 1–29 (cit. on p. 94).

[35] F. Costabile, F. Dell'Accio, and M. I. Gualtieri. "A new approach to Bernoulli polynomials". In: *Rendiconti di Matematica e delle sue Applicazioni. Serie VII* 26.1 (2006), pp. 1–12 (cit. on p. 69).

[36] CharDiales W. Curtis and Irving Reiner. *Representation theory of finite groups and associative algebras*. Vol. 356. American Mathematical Soc., 1966 (cit. on p. 4).

[37] Harm Derksen and Gregor Kemper. *Computational invariant theory*. enlarged. Vol. 130. Encyclopaedia of Mathematical Sciences. With two appendices by Vladimir L. Popov, and an addendum by Norbert A'Campo and Popov, Invariant Theory and Algebraic Transformation Groups, VIII. Springer, Heidelberg, 2015, pp. xxii+366 (cit. on pp. 147–149).

[38] Elena I. Deza. "Cones and polytopes of generalized metrics". In: *Chebyshevskii Sb.* 20.2 (2019), pp. 140–155 (cit. on p. 107).

[39] Persi Diaconis, Susan Holmes, and Mehrdad Shahshahani. "Sampling from a manifold". In: *Advances in modern statistical theory and applications: Festschrift in honor of Morris L. Eaton*. Vol. 10. Inst. Math. Stat. (IMS) Collect. Inst. Math. Statist., Beachwood, OH, 2013, pp. 102–125 (cit. on p. 42).

[40] Persi Diaconis, Gilles Lebeau, and Laurent Michel. "Geometric analysis for the Metropolis algorithm on Lipschitz domains". In: *Invent. Math.* 185.2 (2011), pp. 239–281 (cit. on p. 42).

[41] Persi Diaconis and Laurent Saloff-Coste. "What do we know about the Metropolis algorithm?" In: vol. 57. 1. 27th Annual ACM Symposium on the Theory of Computing (STOC'95) (Las Vegas, NV). 1998, pp. 20–36 (cit. on p. 42).

[42] Fred Diamond and Jerry Shurman. *A first course in modular forms*. Vol. 228. Graduate Texts in Mathematics. Springer-Verlag, New York, 2005, pp. xvi+436 (cit. on p. 64).

[43] Karl Dilcher, Armin Straub, and Christophe Vignat. "Identities for Bernoulli polynomials related to multiple Tornheim zeta functions". In: *Journal of Mathematical Analysis and Applications* 476.2 (2019), pp. 569–584 (cit. on p. 94).

[44] Rick Durrett. *Probability—theory and examples*. Vol. 49. Cambridge Series in Statistical and Probabilistic Mathematics. Fifth edition of [ MR1068527]. Cambridge University Press, Cambridge, 2019, pp. xii+419 (cit. on p. 93).

[45] Alan Edelman and Eric Kostlan. "How many zeros of a random polynomial are real?" In: *American Mathematical Society. Bulletin. New Series* 32.1 (1995), pp. 1–37 (cit. on pp. 129, 132).

[46] Alan Edelman, Eric Kostlan, and Michael Shub. "How many eigenvalues of a random matrix are real?" In: *J. Amer. Math. Soc.* 7.1 (1994), pp. 247–267 (cit. on pp. 129, 132).

[47] David Eisenbud and Joe Harris. *3264 and all that—a second course in algebraic geometry*. Cambridge University Press, Cambridge, 2016, pp. xiv+616 (cit. on p. 46).

[48] Yassine El Maazouz. "The Gaussian entropy map in valued fields". In: *Algebr. Stat.* 13.1 (2022), pp. 1–18 (cit. on pp. 14, 17, 129, 133).

[49] Yassine El Maazouz, Paul Alexander Helminck, and Enis Kaya. "Tropical invariants for binary quintics and reduction types of Picard curves". In: *arXiv:2206.00420* (2022) (cit. on pp. 15, 142).

[50] Yassine El Maazouz and Enis Kaya. "Sampling from $p$-adic algebraic manifolds". In: *arXiv:2207.05911* (2022) (cit. on pp. 14, 41, 129, 133).

[51] Yassine El Maazouz and Antonio Lerario. "Non-archimedean Schur representations of $\mathrm{GL}(n, R)$ and their invariant lattices". In: *arXiv:2209.13634* (2022) (cit. on pp. 15, 128).

[52] Yassine El Maazouz, Gabriele Nebe, and Mima Stanojkovski. "Bolytrope orders". In: *Int. J. Number Theory* (2022). To appear (cit. on pp. 10, 15, 98, 129).

[53] Yassine El Maazouz and Jim Pitman. "The Bernoulli clock: probabilistic and combinatorial interpretations of the Bernoulli polynomials by circular convolution". In: *arXiv:2210.02027* (2022) (cit. on pp. 14, 68).

[54] Yassine El Maazouz and Ngoc M. Tran. "Statistics and tropicalization of local field Gaussian measures". In: *arXiv:1909.00559* (2019) (cit. on pp. 14, 17, 28, 129, 133).

[55] Yassine El Maazouz et al. "Orders and polytropes: matrix algebras from valuations". In: *Beitr. Algebra Geom.* 63.3 (2022), pp. 515–531 (cit. on pp. 10, 15, 98, 102, 129).

[56] Noam Elkies and Abhinav Kumar. "K3 surfaces and equations for Hilbert modular surfaces". In: *Algebra Number Theory* 8.10 (2014), pp. 2297–2411 (cit. on pp. 65, 66).

[57] Antonio J Engler and Alexander Prestel. *Valued fields*. Springer Science & Business Media, 2005 (cit. on pp. 1, 19).

[58] R. C. Entringer. "A combinatorial interpretation of the Euler and Bernoulli numbers". In: *Nieuw Archief voor Wiskunde. Derde Serie* 14 (1966), pp. 241–246 (cit. on p. 81).

[59] Arthur Erdélyi et al. *Higher transcendental functions. Vol. I*. Based, in part, on notes left by Harry Bateman. McGraw-Hill Book Co., Inc., New York-Toronto-London, 1953, pp. xxvi+302, xvii+396 (cit. on pp. 72, 94).

[60] Steven N. Evans. "$p$-adic white noise, chaos expansions, and stochastic integration". In: *Probability measures on groups and related structures, XI (Oberwolfach, 1994)*. World Sci. Publ., River Edge, NJ, 1995, pp. 102–115 (cit. on pp. 19, 41).

[61] Steven N. Evans. "Continuity properties of Gaussian stochastic processes indexed by a local field". In: *Proc. London Math. Soc. (3)* 56.2 (1988), pp. 380–416 (cit. on pp. 128, 130, 132).

[62] Steven N. Evans. "Elementary divisors and determinants of random matrices over a local field". In: *Stochastic Process. Appl.* 102.1 (2002), pp. 89–102 (cit. on pp. 8, 22, 41, 49).

[63] Steven N. Evans. "Equivalence and perpendicularity of local field Gaussian measures". In: *Seminar on Stochastic Processes, 1990 (Vancouver, BC, 1990)*. Vol. 24. Progr. Probab. Birkhäuser Boston, Boston, MA, 1991, pp. 173–181 (cit. on pp. 128, 130, 132).

[64] Steven N. Evans. "Local field Brownian motion". In: *J. Theoret. Probab.* 6.4 (1993), pp. 817–850 (cit. on pp. 41, 128, 130, 132).

[65] Steven N. Evans. "Local fields, Gaussian measures, and Brownian motions". In: *Topics in probability and Lie groups: boundary theory*. Vol. 28. CRM Proc. Lecture Notes. Amer. Math. Soc., Providence, RI, 2001, pp. 11–50 (cit. on pp. 5, 6, 8–10, 17, 19, 27, 28, 31, 40, 41, 128, 130, 132, 134, 140).

[66] Steven N. Evans. "Sample path properties of Gaussian stochastic processes indexed by a local field". In: *Proc. London Math. Soc. (3)* 56.3 (1988), pp. 580–624 (cit. on pp. 128, 130, 132).

[67] Steven N. Evans. "The expected number of zeros of a random system of $p$-adic polynomials". In: *Electron. Comm. Probab.* 11 (2006), pp. 278–290 (cit. on pp. 19, 133).

[68] Steven N. Evans and Tye Lidman. "Expectation, conditional expectation and martingales in local fields". In: *Electronic Journal of Probability* 12.17 (2007), pp. 498–515 (cit. on pp. 128, 130, 132).

[69] Steven N. Evans, Daniel Raban, et al. "Rotatable random sequences in local fields". In: *Electronic Communications in Probability* 24 (2019), Paper No. 37, 12 (cit. on pp. 21, 130, 134).

[70] Gerd Faltings. "Toroidal resolutions for some matrix singularities". In: *Moduli of abelian varieties (Texel Island, 1999)*. Vol. 195. Progr. Math. Birkhäuser, Basel, 2001, pp. 157–184 (cit. on p. 13).

[71] William Fulton. *Young tableaux*. Vol. 35. London Mathematical Society Student Texts. With applications to representation theory and geometry. Cambridge University Press, Cambridge, 1997, pp. x+260 (cit. on p. 135).

[72] William Fulton and Joe Harris. *Representation theory*. Vol. 129. Graduate Texts in Mathematics. A first course, Readings in Mathematics. Springer-Verlag, New York, 1991, pp. xvi+551 (cit. on p. 135).

[73] Yan V. Fyodorov, Antonio Lerario, and Erik Lundberg. "On the number of connected components of random algebraic hypersurfaces". In: *J. Geom. Phys.* 95 (2015), pp. 1–20 (cit. on p. 132).

[74] Damien Gayet and Jean-Yves Welschinger. "Betti numbers of random real hypersurfaces and determinants of random symmetric matrices". In: *J. Eur. Math. Soc. (JEMS)* 18.4 (2016), pp. 733–772 (cit. on pp. 129, 132).

[75] Damien Gayet and Jean-Yves Welschinger. "Expected topology of random real algebraic submanifolds". In: *J. Inst. Math. Jussieu* 14.4 (2015), pp. 673–702 (cit. on pp. 129, 132).

[76] Damien Gayet and Jean-Yves Welschinger. "Lower estimates for the expected Betti numbers of random real hypersurfaces". In: *J. Lond. Math. Soc.* 90 (1 2014), pp. 105–120 (cit. on pp. 129, 132).

[77] Gerard van der Geer. *Hilbert modular surfaces*. Vol. 16. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, 1988, pp. x+291 (cit. on p. 65).

[78] Gerard van der Geer. *Siegel modular forms of degree two and three and invariant theory*. 2021 (cit. on p. 146).

[79] Michel X Goemans. "Semidefinite programming in combinatorial optimization". In: *Mathematical Programming* 79.1-3 (1997), pp. 143–161 (cit. on p. 17).

[80] Roe Goodman and Nolan R. Wallach. *Representations and invariants of the classical groups*. Vol. 68. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1998, pp. xvi+685 (cit. on p. 137).

[81] Paul Gordan. "Beweis, dass jede Covariante und Invariante einer binären Form eine ganze Function mit numerischen Coefficienten einer endlichen Anzahl solcher Formen ist". In: *J. Reine Angew. Math.* 69 (1868), pp. 323–354 (cit. on p. 149).

[82] Ron Graham and Nan Zang. "Enumerating split-pair arrangements". In: *Journal of Combinatorial Theory. Series A* 115.2 (2008), pp. 293–303 (cit. on p. 81).

[83] Caroline Gruson and Vera Serganova. *A journey through representation theory*. Universitext. Springer, Cham, 2018 (cit. on p. 135).

[84] Walter Gubler. "A guide to tropicalizations". In: *Algebraic and combinatorial aspects of tropical geometry*. Vol. 589. Contemp. Math. Amer. Math. Soc., Providence, RI, 2013, pp. 125–189 (cit. on p. 154).

[85] Marvin Anas Hahn and Binglin Li. "Mustafin varieties, moduli spaces and tropical geometry". In: *Manuscripta Math.* 166.1-2 (2021), pp. 167–197 (cit. on p. 13).

[86] Joe Harris. *Algebraic geometry*. Vol. 133. Graduate Texts in Mathematics. A first course, Corrected reprint of the 1992 original. Springer-Verlag, New York, 1995, pp. xx+328 (cit. on p. 46).

[87] Paul Alexander Helminck. "Invariants for trees of non-archimedean polynomials and skeleta of superelliptic curves". In: *Mathematische Zeitschrift* (2022) (cit. on pp. 144, 146, 166).

[88] Paul Alexander Helminck. *Skeletal filtrations of the fundamental group of a non-archimedean curve*. 2021 (cit. on p. 166).

[89] Paul Alexander Helminck. *Tropical Igusa Invariants*. 2016 (cit. on pp. 66, 143, 146, 159).

[90] Friedrich Hirzebruch. *Topological methods in algebraic geometry*. Classics in Mathematics. Translated from the German and Appendix One by R. L. E. Schwarzenberger, With a preface to the third English edition by the author and Schwarzenberger, Appendix Two by A. Borel, Reprint of the 1978 edition. Springer-Verlag, Berlin, 1995, pp. xii+234 (cit. on p. 69).

[91] Joseph D. Horton and Andrew Kurn. "Counting sequences with complete increasing subsequences". In: *Congr. Numer.* 33 (1981), pp. 75–80 (cit. on pp. 2, 14, 71, 88, 96).

[92] Nobuyuki Ikeda and Setsuo Taniguchi. "Euler polynomials, Bernoulli polynomials, and Lévy's stochastic area formula". In: *Bull. Sci. Math.* 135.6-7 (2011), pp. 684–694 (cit. on p. 69).

[93] Nobuyuki Ikeda and Setsuo Taniguchi. "The Itô-Nisio theorem, quadratic Wiener functionals, and 1-solitons". In: *Stochastic Process. Appl.* 120.5 (2010), pp. 605–621 (cit. on p. 69).

[94] Lizhen Ji. "From symmetric spaces to buildings, curve complexes and outer spaces". In: *Innov. Incidence Geom.* 10 (2009), pp. 33–80 (cit. on pp. 10, 129).

[95] Charles Jordan. *Calculus of finite differences*. Third. Introduction by Harry C. Carver. Chelsea Publishing Co., New York, 1965, pp. xxi+655 (cit. on p. 95).

[96] Michael Joswig. "Essentials of tropical combinatorics". In: *Book in preparation*. Graduate Studies in Mathematics 1 (2014), p. 226 (cit. on pp. 13, 99, 100, 104–107, 115).

[97] Michael Joswig and Katja Kulas. "Tropical and ordinary convexity combined". In: *Advances in Geometry* 10.2 (2010), pp. 333–352 (cit. on pp. 104, 106, 107, 109, 112).

[98] Mark Kac. "On a characterization of the normal distribution". In: *American Journal of Mathematics* 61.3 (1939), pp. 726–728 (cit. on p. 9).

[99] Kiumars Kaveh and Christopher Manon. "Gröbner theory and tropical geometry on spherical varieties". In: *Transform. Groups* 24.4 (2019), pp. 1095–1145 (cit. on p. 146).

[100] Neal Koblitz. *p-adic numbers, p-adic analysis, and zeta-functions*. Second. Vol. 58. Graduate Texts in Mathematics. Springer-Verlag, New York, 1984, pp. xii+150 (cit. on p. 19).

[101] Eric Kostlan. "On the distribution of roots of random polynomials". In: *From Topology to Computation: Proceedings of the Smalefest (Berkeley, CA, 1990)*. Springer, New York, 1993, pp. 419–431 (cit. on pp. 130, 132).

[102] Eric Kostlan. "On the expected number of real roots of a system of random polynomial equations". In: *Foundations of computational mathematics (Hong Kong, 2000)*. World Sci. Publ., River Edge, NJ, 2002, pp. 149–188 (cit. on pp. 129, 132).

[103] DM Koteljanskii. "A property of sign-symmetric matrices". In: *Amer. Math. Soc. Transl. Ser* 2.27 (1963), pp. 19–23 (cit. on p. 18).

[104] Jeroen Kuipers, Dries Vermeulen, and Mark Voorneveld. "A generalization of the Shapley-Ichiishi result". In: *Internat. J. Game Theory* 39.4 (2010), pp. 585–602 (cit. on p. 26).

[105] Avinash Kulkarni and Antonio Lerario. "$p$-adic integral geometry". In: *SIAM J. Appl. Algebra Geom.* 5.1 (2021), pp. 28–59 (cit. on pp. 19, 41, 43, 47, 55, 129, 133, 134).

[106] Joseph Louis Lagrange. *Oeuvres de Lagrange*. Vol. 13. Gauthier-Villars, 1882 (cit. on p. 80).

[107] Derrick Henry Lehmer. "A new approach to Bernoulli polynomials". In: *American Mathematical Monthly* 95.10 (1988), pp. 905–911 (cit. on p. 69).

[108] Derrick Henry Lehmer. "On the maxima and minima of Bernoulli polynomials". In: *American Mathematical Monthly* 47 (1940), pp. 533–538 (cit. on p. 69).

[109] Antonio Lerario. "Random matrices and the average topology of the intersection of two quadrics". In: *Proc. Amer. Math. Soc.* 143.8 (2015), pp. 3239–3251 (cit. on pp. 129, 132).

[110] Antonio Lerario and Erik Lundberg. "Gap probabilities and Betti numbers of a random intersection of quadrics". In: *Discrete Comput. Geom.* 55.2 (2016), pp. 462–496 (cit. on pp. 129, 132).

[111] Paul Lévy. "Le mouvement brownien plan". In: *Amer. J. Math.* 62 (1940), pp. 487–550 (cit. on p. 69).

[112] Paul Lévy. "Wiener's random function, and other Laplacian random functions". In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. University of California Press, Berkeley-Los Angeles, Calif., 1951, pp. 171–187 (cit. on p. 69).

[113] Bo Lin, Anthea Monod, and Ruriko Yoshida. "Tropical foundations for probability & statistics on phylogenetic tree space". In: *Calhoun: The NPS Institutional Archive* (2018) (cit. on p. 19).

[114] Jun S. Liu. *Monte Carlo strategies in scientific computing*. Springer Series in Statistics. Springer-Verlag, New York, 2001, pp. xvi+343 (cit. on p. 42).

[115] Qing Liu. "Courbes stables de genre 2 et leur schéma de modules". In: *Math. Ann.* 295.2 (1993), pp. 201–222 (cit. on pp. 143, 146).

[116] Diane Maclagan and Bernd Sturmfels. *Introduction to tropical geometry*. Vol. 161. Graduate Studies in Mathematics. American Mathematical Soc., 2015, pp. xii+363 (cit. on pp. 100, 102, 104, 108, 155).

[117] Wilhelm Magnus. "On the exponential solution of differential equations for a linear operator". In: *Comm. Pure Appl. Math.* 7 (1954), pp. 649–673 (cit. on p. 69).

[118] Rida Ait El Manssour and Antonio Lerario. "Probabilistic enumerative geometry over $p$-adic numbers: linear spaces on complete intersections". In: *arXiv:2011.07558* (2020) (cit. on p. 41).

[119] Hannah Markwig. "Tropical curves and covers and their moduli spaces". In: *Jahresber. Dtsch. Math.-Ver.* 122.3 (2020), pp. 139–166 (cit. on p. 157).

[120] Nathaniel F. G. Martin and James W. England. *Mathematical theory of entropy*. Vol. 12. Encyclopedia of Mathematics and its Applications. With a foreword by James K. Brooks. Addison-Wesley Publishing Co., Reading, Mass., 1981, pp. xxi+257 (cit. on p. 18).

[121] Barry Mazur. "How can we construct abelian Galois extensions of basic number fields?" In: *Bull. Amer. Math. Soc. (N.S.)* 48.2 (2011), pp. 155–209 (cit. on p. 69).

[122] Mateusz Michałek and Bernd Sturmfels. "Invitation to nonlinear algebra". In: *Graduate Studies in Mathematics, American Mathematical Society* (2019) (cit. on pp. 17, 136).

[123] James S. Milne. *Algebraic groups*. Vol. 170. Cambridge Studies in Advanced Mathematics. The theory of group schemes of finite type over a field. Cambridge University Press, Cambridge, 2017, pp. xvi+644 (cit. on p. 63).

[124] John W. Milnor and Michel A. Kervaire. "Bernoulli numbers, homotopy groups, and a theorem of Rohlin". In: *Proc. Internat. Congress Math. 1958*. Cambridge Univ. Press, New York, 1960, pp. 454–458 (cit. on p. 69).

[125] Hugh L. Montgomery. *Early Fourier analysis*. Vol. 22. Pure and Applied Undergraduate Texts. American Mathematical Society, Providence, RI, 2014, pp. x+390 (cit. on p. 69).

[126] Louis Joel Mordell. "Expansion of a function in a series of Bernoulli polynomials, and some other polynomials". In: *Journal of Mathematical Analysis and Applications* 15 (1966), pp. 132–140 (cit. on p. 95).

[127] David Mumford, John Fogarty, and Frances Clare Kirwan. *Geometric invariant theory*. Third. Vol. 34. Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)]. Springer-Verlag, Berlin, 1994, pp. xiv+292 (cit. on p. 147).

[128] Fedor Nazarov and Mikhail Sodin. "On the number of nodal domains of random spherical harmonics". In: *Amer. J. Math.* 131.5 (2009), pp. 1337–1357 (cit. on pp. 129, 132).

[129] Gabriele Nebe and Allan Steel. "Recognition of division algebras". In: *J. Algebra* 322.3 (2009), pp. 903–909 (cit. on p. 119).

[130] Niels Erik Nörlund. *Vorlesungen über Differenzenrechnung*. Vol. 13. J. Springer, 1924 (cit. on pp. 70, 94).

[131] Sam Payne. "Analytification is the limit of all tropicalizations". In: *Math. Res. Lett.* 16.3 (2009), pp. 543–556 (cit. on p. 153).

[132] Jim Pitman and Marc Yor. "Infinitely divisible laws associated with hyperbolic functions". In: *Canadian Journal of Mathematics. Journal Canadien de Mathématiques* 55.2 (2003), pp. 292–330 (cit. on p. 69).

[133] Wilhelm Plesken. *Group rings of finite groups over p-adic integers*. Vol. 1026. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1983, pp. ii+151 (cit. on pp. 15, 99–102, 106, 108, 110).

[134] Mihnea Popa. "Chapter 3: $p$-adic integration". In: *https://people.math.harvard.edu/ mpopa/571/* (2011). Unpublished notes (cit. on p. 47).

[135] Mihaela Ileana Popoviciu Draisma. "Invariants of binary forms". en. In: (2014) (cit. on pp. 147, 149–151).

[136] Gopal Prasad. "Finite group actions on reductive groups and buildings and tamely-ramified descent in Bruhat-Tits theory". In: *Amer. J. Math.* 142.4 (2020), pp. 1239–1267 (cit. on pp. 10, 129).

[137] Gopal Prasad and Jiu-Kang Yu. "On finite group actions on reductive groups and buildings". In: *Invent. Math.* 147.3 (2002), pp. 545–560 (cit. on pp. 10, 129).

[138] Irving Reiner. *Maximal orders*. London Mathematical Society Monographs, No. 5. Academic Press [Harcourt Brace Jovanovich, Publishers], London-New York, 1975, pp. xii+395 (cit. on p. 119).

[139] Dan Romik. "On the number of $n$-dimensional representations of SU(3), the Bernoulli numbers, and the Witten zeta function". In: *Acta Arith.* 180.2 (2017), pp. 111–159 (cit. on p. 69).

[140] Arnoud C. M. van Rooij. *Non-Archimedean functional analysis*. Vol. 51. Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 1978, pp. x+404 (cit. on pp. 4, 6, 19).

[141] Guy Rousseau. "Euclidean buildings". In: *Géométries à courbure négative ou nulle, groupes discrets et rigidités*. Vol. 18. Sémin. Congr. Soc. Math. France, Paris, 2009, pp. 77–116 (cit. on pp. 10, 129).

[142] The Sage Developers. "SageMath, the Sage Mathematics Software System (Version 9.5)". In: (2022). `https://www.sagemath.org` (cit. on pp. 46, 147).

[143] Peter Sarnak. "Letter to B. Gross and J. Harris on ovals of random planes curve". In: (2011). available at `http://publications.poedu/sarnak/section/515` (cit. on pp. 129, 132).

[144] Wilhelmus Hendricus Schikhof. *Ultrametric Calculus: an introduction to p-adic analysis*. Vol. 4. Cambridge University Press, 2007 (cit. on pp. 1, 4, 19).

[145] Peter Schneider. *p-adic Lie groups*. Vol. 344. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Heidelberg, 2011, pp. xii+254 (cit. on p. 47).

[146] Jean-Pierre Serre. *Local fields*. Vol. 67. Graduate Texts in Mathematics. Translated from the French by Marvin Jay Greenberg. Springer-Verlag, New York-Berlin, 1979, pp. viii+241 (cit. on pp. 1, 19, 42).

[147] Jean-Pierre Serre. "Quelques applications du théorème de densité de Chebotarev". In: *Inst. Hautes Études Sci. Publ. Math.* 5.54 (1981), pp. 323–401 (cit. on pp. 43, 55).

[148] Jean-Pierre Serre. *Trees*. Translated from the French by John Stillwell. Springer-Verlag, Berlin-New York, 1980, pp. ix+142 (cit. on p. 125).

[149] Michael Shub and Steve Smale. "Complexity of Bézout's theorem. I. Geometric aspects". In: *J. Amer. Math. Soc.* 6.2 (1993), pp. 459–501 (cit. on pp. 129, 132).

[150] Michael Shub and Steve Smale. "Complexity of Bezout's theorem. II. Volumes and probabilities". In: *Computational algebraic geometry (Nice, 1992)*. Ed. by F. Eyssette and A. Galligo. Vol. 109. Progr. Math. Birkhäuser Boston, Boston, MA, 1993, pp. 267–285 (cit. on pp. 129, 132).

[151] Michael Shub and Steve Smale. "Complexity of Bezout's theorem. III. Condition number and packing". In: *J. Complexity* 9.1 (1993). Festschrift for Joseph F. Traub, Part I, pp. 4–14 (cit. on pp. 129, 132).

[152] Joseph H. Silverman. *The arithmetic of elliptic curves*. Second. Vol. 106. Graduate Texts in Mathematics. Springer, Dordrecht, 2009, pp. xx+513 (cit. on pp. 64, 143, 159).

[153] Richard P. Stanley. *Enumerative combinatorics. Volume 1.* Second. Vol. 49. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2012, pp. xiv+626 (cit. on pp. 81, 88).

[154] Bernd Sturmfels. *Algorithms in invariant theory.* Second. Texts and Monographs in Symbolic Computation. SpringerWienNewYork, Vienna, 2008, pp. vi+197 (cit. on pp. 142, 147).

[155] Bernd Sturmfels. "Open problems in algebraic statistics". In: *Emerging applications of algebraic geometry.* Springer, 2009, pp. 351–363 (cit. on pp. 27, 28).

[156] Bernd Sturmfels and Caroline Uhler. "Multivariate Gaussian, semidefinite matrix completion, and convex algebraic geometry". In: *Ann. Inst. Statist. Math.* 62.4 (2010), pp. 603–638 (cit. on p. 17).

[157] Ping Sun. "Moment representation of Bernoulli polynomial, Euler polynomial and Gegenbauer polynomials". In: *Statistics & Probability Letters* 77.7 (2007), pp. 748–751 (cit. on p. 69).

[158] Andrew V. Sutherland. "Constructing elliptic curves over finite fields with prescribed torsion". In: *Math. Comp.* 81.278 (2012), pp. 1131–1147 (cit. on p. 64).

[159] Mitchell H. Taibleson. *Fourier analysis on local fields.* Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1975, pp. xii+294 (cit. on p. 5).

[160] Patrice Tauvel and Rupert W. T. Yu. *Lie algebras and algebraic groups.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005, pp. xvi+653 (cit. on p. 137).

[161] Ngoc M. Tran. "Enumerating polytropes". In: *Journal of Combinatorial Theory, Series A* 151 (2017), pp. 1–22 (cit. on pp. 100, 106).

[162] Ngoc M. Tran. "Tropical Gaussians: a brief survey". In: *Algebraic statistics* 11 (2020) (cit. on p. 19).

[163] Ngoc M. Tran and Josephine Yu. "Product-mix auctions and tropical geometry". In: *Math. Oper. Res.* 44.4 (2019), pp. 1396–1411 (cit. on p. 19).

[164] Fang-Ting Tu. "On orders of $M(2, K)$ over a non-Archimedean local field". In: *Int. J. Number Theory* 7.5 (2011), pp. 1137–1149 (cit. on pp. 125, 127).

[165] Vasilii Sergeevich Vladimirov, Igor' Vasil'evich Volovich, and Evgenii Igorevich Zelenov. *p-adic analysis and mathematical physics.* Vol. 1. Series on Soviet and East European Mathematics. World Scientific Publishing Co., Inc., River Edge, NJ, 1994, pp. xx+319 (cit. on p. 41).

[166] André Weil. *Basic number theory.* Vol. 144. Springer Science & Business Media, 2013 (cit. on pp. 6, 19).

[167] Ruriko Yoshida, Leon Zhang, and Xu Zhang. "Tropical principal component analysis and its application to phylogenetics". In: *Bull. Math. Biol.* 81.2 (2019), pp. 568–597 (cit. on p. 19).

[168] Stephen M. Zemyan. "On the zeroes of the $N$th partial sum of the exponential series". In: *Amer. Math. Monthly* 112.10 (2005), pp. 891–909 (cit. on pp. 71, 90, 92).

[169] Leon Zhang. "Computing min-convex hulls in the affine building of $\mathrm{SL}_d$". In: *Discrete Comput. Geom.* 65.4 (2021), pp. 1314–1336 (cit. on pp. 13, 112, 113).