

Relation Lattices and Kolmogorov Deficiency of Rational Points

Luca Blachi

June 2026

Abstract

Let $v \in \mathbb{Z}_{\text{prim}}^{d+1}$ be a primitive integer vector, regarded as a rational point $[v] \in \mathbb{P}^d(\mathbb{Q})$. We study the lattice of homogeneous integer linear relations

$$\Lambda_v = \{a \in \mathbb{Z}^{d+1} : a \cdot v = 0\}$$

and the affine lattice of Bezout certificates

$$\mathcal{B}_v = \{w \in \mathbb{Z}^{d+1} : v \cdot w = 1\}.$$

The lattice Λ_v has rank d and covolume $|v|_2$, so the natural scale for both short relations and short Bezout certificates is $|v|_2^{1/d}$.

We first prove an elementary sharp threshold theorem: for almost all primitive vectors $v \in \mathbb{Z}^{d+1}$ with $|v|_2 \leq R$,

$$\lambda_1(\Lambda_v) = R^{1/d+o(1)}.$$

Equivalently, a typical rational point of height R satisfies no rational hyperplane relation of height $< R^{1/d-o(1)}$, while every point satisfies one of height $O_d(R^{1/d})$.

We then use the equidistribution theorem of Horesh and Karasik for primitive vectors, orthogonal lattices, and shortest solutions to the gcd equation. Their theorem implies that the normalized relation lattices

$$\widehat{\Lambda}_v = |v|_2^{-1/d} \Lambda_v$$

equidistribute in the corresponding space of unimodular Euclidean lattice shapes. Consequently, the successive minima of Λ_v , after normalization by $|v|_2^{1/d}$, have the limiting law of the successive minima of a Haar-random unimodular d -lattice.

Finally, we translate these facts into finite-height Kolmogorov deficiency. For an integer $R \geq 1$, set

$$V_{d+1}(R) = \{v \in \mathbb{Z}_{\text{prim}}^{d+1} : 0 < |v|_2 \leq R\}$$

and

$$\delta(v; R) = \log_2 |V_{d+1}(R)| - K(v \mid d, R).$$

If

$$\lambda_1(\Lambda_v) \leq t|v|_2^{1/d}$$

for $t \in \mathbb{Q}_{>0}$, then at continuity points of the limiting distribution,

$$\delta(v; R) \geq -\log_2 F_d(t) - K(t) - O_d(1) - o_R(1),$$

where $F_d(t)$ is the Haar probability that a unimodular d -lattice has a nonzero vector of length at most t . For $d \geq 2$,

$$F_d(t) \sim \frac{\text{vol}(B_d)}{2\zeta(d)} t^d \quad (t \downarrow 0),$$

so an anomalously short relation forces approximately $d \log_2(1/t)$ bits of Kolmogorov deficiency, up to the cost of describing t . We also record the corresponding Bezout-certificate consequence of the Horesh-Karasik shortest-solution theorem.

1 Introduction

A rational point

$$x = [x_0 : \cdots : x_d] \in \mathbb{P}^d(\mathbb{Q})$$

may be represented by a primitive vector

$$v = (x_0, \dots, x_d) \in \mathbb{Z}_{\text{prim}}^{d+1},$$

unique up to sign. A rational hyperplane through x is given by a nonzero vector

$$a = (a_0, \dots, a_d) \in \mathbb{Z}^{d+1}$$

such that

$$a \cdot v = 0.$$

Thus the set of all rational hyperplane relations through x is encoded by the lattice

$$\Lambda_v = v^\perp \cap \mathbb{Z}^{d+1}.$$

This paper studies Λ_v from three viewpoints.

First, Λ_v is a geometry-of-numbers object. It has rank d and covolume $|v|_2$, so one expects its shortest nonzero vector to have length about $|v|_2^{1/d}$. We prove the almost-everywhere threshold

$$\lambda_1(\Lambda_v) = R^{1/d+o(1)}$$

for primitive vectors v with $|v|_2 \leq R$. This is an elementary result obtained from Minkowski's theorem and a counting argument.

Second, Λ_v has a limiting random-lattice profile. The deep input is a theorem of Horesh and Karasik, who prove joint equidistribution of primitive vectors, their orthogonal lattices, and shortest solutions to the gcd equation. We use their result as an external theorem. It implies that the normalized relation lattices

$$\widehat{\Lambda}_v = |v|_2^{-1/d} \Lambda_v$$

equidistribute in the natural homogeneous space of unimodular lattice shapes. Hence successive minima, covering radius, and other orthogonally invariant lattice-shape statistics of Λ_v have limiting distributions.

Third, Λ_v has an algorithmic interpretation. A short vector in Λ_v is a short linear explanation of the rational point $[v]$. We define the finite-height Kolmogorov deficiency

$$\delta(v; R) = \log_2 |V_{d+1}(R)| - K(v \mid d, R),$$

where

$$V_{d+1}(R) = \{v \in \mathbb{Z}_{\text{prim}}^{d+1} : 0 < |v|_2 \leq R\}$$

and $R \in \mathbb{N}$. A vector has small deficiency if it is algorithmically typical inside the height ball. We show that cusp events of the relation lattice, such as having an anomalously short relation, force Kolmogorov deficiency. This is an arithmetic instance of the two-part coding principle from algorithmic statistics: membership in a small, effectively specified model gives a shorter description.

Finally, for primitive v , the equation

$$v \cdot w = 1$$

has integer solutions. Such a w is a Bezout certificate that

$$\gcd(v_0, \dots, v_d) = 1.$$

The length of the shortest such certificate is governed by the same scale $|v|_2^{1/d}$, and Horesh-Karasik's theorem gives its limiting distribution after the appropriate normalization. Thus linear explanations and certificates of primitivity are controlled by the same random-lattice profile.

The contribution of this paper is not a new homogeneous-dynamics equidistribution theorem. That input is due to Horesh and Karasik. Rather, we isolate the relation-lattice profile of rational points, prove the elementary sharp threshold for the shortest relation, and derive finite-height Kolmogorov-deficiency consequences from the random-lattice limit.

2 Primitive vectors and relation lattices

Fix $d \geq 1$, and put

$$N = d + 1.$$

For an integer $R \geq 1$, define

$$V_N(R) = \{v \in \mathbb{Z}_{\text{prim}}^N : 0 < |v|_2 \leq R\}.$$

It is classical that

$$|V_N(R)| \sim \frac{\text{vol}(B_N)}{\zeta(N)} R^N.$$

For $v \in \mathbb{Z}_{\text{prim}}^N$, define the relation lattice

$$\Lambda_v = \{a \in \mathbb{Z}^N : a \cdot v = 0\}.$$

It is a lattice of rank $N - 1 = d$ in the Euclidean hyperplane $v^\perp$.

Lemma 2.1. For primitive $v \in \mathbb{Z}^N$,

$$\text{covol}(\Lambda_v) = |v|_2.$$

Proof.

The homomorphism

$$\varphi_v : \mathbb{Z}^N \rightarrow \mathbb{Z}, \quad a \mapsto a \cdot v,$$

is surjective because v is primitive. Its kernel is Λ_v . The covolume of this kernel in $v^\perp$ is the Euclidean norm of the primitive normal vector v . Hence

$$\text{covol}(\Lambda_v) = |v|_2.$$

□

Define the normalized relation lattice

$$\widehat{\Lambda}_v = |v|_2^{-1/d} \Lambda_v.$$

This has covolume 1. Since it lives in $v^\perp$, one must choose an isometry

$$v^\perp \simeq \mathbb{R}^d$$

to view it as a unimodular lattice in $\mathbb{R}^d$. The resulting lattice shape is independent of this choice up to orthogonal transformation, and all statistics used in this paper are orthogonally invariant. Let

$$\lambda_1(v) \leq \dots \leq \lambda_d(v)$$

be the successive minima of Λ_v with respect to the Euclidean norm.

3 The elementary threshold for the shortest relation

The first result is independent of Horesh-Karasik.

Theorem 3.1. For every $\varepsilon > 0$,

$$\frac{|\{v \in V_N(R) : \lambda_1(v) \leq R^{1/d-\varepsilon}\}|}{|V_N(R)|} \rightarrow 0.$$

Moreover, for every $v \in V_N(R)$,

$$\lambda_1(v) \ll_d R^{1/d}.$$

Consequently, for almost all $v \in V_N(R)$,

$$\lambda_1(v) = R^{1/d+o(1)}.$$

Proof.

The upper bound follows from Minkowski's theorem. By Lemma 2.1,

$$\text{covol}(\Lambda_v) = |v|_2 \leq R.$$

Since Λ_v has rank d , Minkowski's theorem gives

$$\lambda_1(v) \ll_d R^{1/d}.$$

For the lower bound, let $A \leq R$. We count primitive $v \in V_N(R)$ for which there exists a nonzero $a \in \mathbb{Z}^N$ with

$$|a|_2 \leq A \quad \text{and} \quad a \cdot v = 0.$$

For fixed primitive a , the lattice

$$a^\perp \cap \mathbb{Z}^N$$

has rank d and covolume $|a|_2$. By the standard Lipschitz principle applied to this lattice,

$$|\{v \in \mathbb{Z}^N : |v|_2 \leq R, a \cdot v = 0\}| \ll_d \frac{R^d}{|a|_2} + R^{d-1}.$$

Indeed, the main term is the volume of the d -dimensional section divided by the covolume $|a|_2$, and the boundary contribution is $O_d(R^{d-1})$.

Summing over primitive a with $0 < |a|_2 \leq A$, and using that the number of integer vectors with Euclidean norm in $[r, r+1]$ is $O_d(r^d)$, we get

$$\ll_d R^d \sum_{r \leq A} r^{d-1} + R^{d-1} A^{d+1} \ll_d R^d A^d + R^{d-1} A^{d+1}.$$

For $A \leq R$, this is

$$\ll_d R^d A^d.$$

Since

$$|V_N(R)| \asymp_d R^{d+1},$$

we obtain

$$\mathbb{P}(\lambda_1(v) \leq A) \ll_d \frac{A^d}{R}.$$

Taking

$$A = R^{1/d-\varepsilon}$$

gives

$$\mathbb{P}(\lambda_1(v) \leq R^{1/d-\varepsilon}) \ll_d R^{-d\varepsilon},$$

which tends to 0. $\square$

Thus the exponent $1/d$ is sharp for the shortest rational hyperplane relation.

4 Orthogonal-lattice equidistribution

Let $\mathcal{L}_d$ denote the space of unimodular Euclidean lattice shapes of rank d , equipped with its Haar probability measure μ_d . We use the following consequence of the theorem of Horesh and Karasik.

Theorem 4.1. Horesh-Karasik input. As $R \rightarrow \infty$, the normalized orthogonal lattices

$$\widehat{\Lambda}_v = |v|_2^{-1/d} \Lambda_v, \quad v \in V_N(R),$$

equidistribute in $\mathcal{L}_d$. Equivalently, for every bounded continuous orthogonally invariant function F on the space of unimodular d -lattices,

$$\frac{1}{|V_N(R)|} \sum_{v \in V_N(R)} F(\widehat{\Lambda}_v) \longrightarrow \int_{\mathcal{L}_d} F(\Lambda) d\mu_d(\Lambda).$$

Horesh and Karasik prove a stronger joint equidistribution theorem involving the direction of v , the orthogonal lattice, and shortest solutions to the gcd equation. We use only the lattice-shape and shortest-solution consequences.

Corollary 4.2. Successive minima law. At every continuity point of the limiting distribution,

$$\left(\frac{\lambda_1(v)}{|v|_2^{1/d}}, \dots, \frac{\lambda_d(v)}{|v|_2^{1/d}} \right)$$

converges in distribution to

$$(\lambda_1(\Lambda), \dots, \lambda_d(\Lambda)),$$

where Λ is a Haar-random unimodular d -lattice.

In particular, for every $\varepsilon > 0$,

$$\lambda_i(v) = |v|_2^{1/d+o(1)} \quad (1 \leq i \leq d)$$

for almost all $v \in V_N(R)$.

Proof.

The successive minima are measurable and continuous outside a μ_d -null set. The assertion follows from Theorem 4.1 and the continuous mapping theorem. $\square$

This gives the full independent-relation profile of a typical rational point.

5 The shortest-relation distribution

Let

$$F_d(t) = \mu_d\{\Lambda \in \mathcal{L}_d : \lambda_1(\Lambda) \leq t\}.$$

Corollary 5.1. For every continuity point $t > 0$ of F_d ,

$$\lim_{R \rightarrow \infty} \frac{|\{v \in V_N(R) : \lambda_1(v) \leq t|v|_2^{1/d}\}|}{|V_N(R)|} = F_d(t).$$

For $d \geq 2$,

$$F_d(t) \sim \frac{\text{vol}(B_d)}{2\zeta(d)} t^d \quad (t \downarrow 0).$$

Proof.

The convergence statement is the first-minimum case of Corollary 4.2.

For the small- t asymptotic, use the standard cusp estimate for the space of unimodular lattices. By Siegel's mean value theorem, the expected number of primitive lattice vectors in the Euclidean ball of radius t is

$$\frac{\text{vol}(B_d)}{\zeta(d)} t^d.$$

For $t \downarrow 0$, Rogers' second moment estimate implies that the probability of two independent primitive vectors in the ball is $O_d(t^{2d})$. Since primitive vectors occur in pairs $\pm u$, the probability of a nonzero vector of length at most t is therefore

$$\frac{\text{vol}(B_d)}{2\zeta(d)} t^d + o(t^d).$$

□

6 Kolmogorov deficiency and short relations

Fix a universal prefix-free Turing machine. For integer $R \geq 1$, define

$$\delta(v; R) = \log_2 |V_N(R)| - K(v \mid N, R).$$

We first record the precise coding lemma used below.

Lemma 6.1. Coding lemma. Let $(E_R)_{R \in \mathbb{N}}$ be a uniformly decidable family of finite subsets

$$E_R \subseteq V_N(R).$$

Then for every $v \in E_R$,

$$K(v \mid N, R, \mathcal{E}) \leq \lceil \log_2 |E_R| \rceil + O(1),$$

where $\mathcal{E}$ denotes the fixed algorithm deciding and enumerating E_R .

If $E_R = E_R(\tau)$ depends on a finite parameter τ , then

$$K(v \mid N, R) \leq K(\tau) + \lceil \log_2 |E_R(\tau)| \rceil + O_N(1).$$

Proof.

Given N, R , the enumeration algorithm, and the index of v in a canonical enumeration of E_R , the decoder recovers v . The index costs

$$\lceil \log_2 |E_R| \rceil + O(1)$$

bits. If the family depends on a parameter τ , include a prefix-free description of τ . □

Let $t \in \mathbb{Q}_{>0}$, and define

$$E_R(t) = \{v \in V_N(R) : \lambda_1(v) \leq t|v|_2^{1/d}\}.$$

Since $t \in \mathbb{Q}_{>0}$, the set $E_R(t)$ is uniformly decidable from N, R, t . Indeed, one enumerates all $a \in \mathbb{Z}^N$ satisfying

$$|a|_2 \leq tR^{1/d}$$

and checks whether $a \cdot v = 0$ for some nonzero a . Equivalently, to avoid radicals, compare

$$|a|_2^{2d} \leq t^{2d} R^2$$

after clearing denominators.

Theorem 6.2. Deficiency law for short relations. Let $t \in \mathbb{Q}_{>0}$ be a continuity point of F_d . If

$$v \in E_R(t),$$

then

$$\delta(v; R) \geq -\log_2 F_d(t) - K(t) - O_d(1) - o_R(1).$$

For $d \geq 2$, as $t \downarrow 0$ through positive rationals at continuity points of F_d ,

$$\lambda_1(\Lambda_v) \leq t|v|_2^{1/d}$$

forces

$$\delta(v; R) \geq d \log_2(1/t) - K(t) - O_d(1) - o_t(1) - o_R(1).$$

Proof.

By Lemma 6.1,

$$K(v \mid N, R) \leq K(t) + \log_2 |E_R(t)| + O_d(1).$$

Thus

$$\delta(v; R) \geq -\log_2 \frac{|E_R(t)|}{|V_N(R)|} - K(t) - O_d(1).$$

By Corollary 5.1,

$$\frac{|E_R(t)|}{|V_N(R)|} \rightarrow F_d(t),$$

which gives the first assertion. The small- t estimate follows from

$$F_d(t) \sim \frac{\text{vol}(B_d)}{2\zeta(d)} t^d.$$

□

This theorem says that cusp excursions of the relation lattice are finite-height Kolmogorov deficiency events. The term $K(t)$ is necessary if t is allowed to vary.

7 Bezout certificates

For $v \in \mathbb{Z}_{\text{prim}}^N$, define

$$\mathcal{B}_v = \{w \in \mathbb{Z}^N : v \cdot w = 1\}.$$

This is a translate of Λ_v . Define

$$b(v) = \min\{|w|_2 : w \in \mathcal{B}_v\}.$$

A vector $w \in \mathcal{B}_v$ is a Bezout certificate proving that the coordinates of v are coprime.

Horesh and Karasik prove joint equidistribution of the orthogonal lattice and the shortest solutions to the gcd equation. We record the consequence needed here.

Theorem 7.1. Bezout certificate scale and law. There is a probability distribution ν_d on $[0, \infty)$, determined by the Horesh-Karasik limiting distribution, such that

$$\frac{b(v)}{|v|_2^{1/d}}$$

converges in distribution to ν_d as v ranges over $V_N(R)$ and $R \rightarrow \infty$.

In particular,

$$b(v) = |v|_2^{1/d+o(1)}$$

for almost all $v \in V_N(R)$.

Proof.

This is the shortest-gcd-solution component of the Horesh-Karasik joint equidistribution theorem. The scale $|v|_2^{1/d}$ is the scale of the covering radius of the relation lattice Λ_v , whose covolume is $|v|_2$. $\square$

Corollary 7.2. Deficiency from short Bezout certificates. Let

$$G_d(t) = \nu_d([0, t]).$$

At every continuity point $t \in \mathbb{Q}_{>0}$ of G_d , if

$$b(v) \leq t|v|_2^{1/d},$$

then

$$\delta(v; R) \geq -\log_2 G_d(t) - K(t) - O_d(1) - o_R(1).$$

Proof.

Apply the coding lemma to the uniformly decidable family

$$\{v \in V_N(R) : b(v) \leq t|v|_2^{1/d}\},$$

and use Theorem 7.1 to identify its limiting density. $\square$

Thus unusually short certificates of primitivity are also Kolmogorov deficiency events.

8 Projective height and norm dependence

The paper has used Euclidean norm in order to match the homogeneous-dynamics input. The usual projective height of

$$[v] \in \mathbb{P}^d(\mathbb{Q})$$

is

$$H([v]) = |v|_\infty$$

for a primitive representative v .
Since

$$|v|_\infty \leq |v|_2 \leq \sqrt{N}|v|_\infty,$$

we have the inclusions

$$\{v : |v|_2 \leq B\} \subseteq \{v : |v|_\infty \leq B\} \subseteq \{v : |v|_2 \leq \sqrt{N}B\}.$$

Both Euclidean and sup-norm height balls have cardinality $\asymp_N B^N$. Therefore every density-one statement with exponent $B^{1/d+o(1)}$ transfers between Euclidean height and projective sup-norm height.

Thus, for almost all rational points $x = [v]$ with $H(x) \leq B$, the shortest rational hyperplane through x , measured in any fixed Euclidean or sup norm on coefficient vectors, has height

$$B^{1/d+o(1)}.$$

A full limiting distribution for box height would require a box-height analogue of the orthogonal-lattice equidistribution theorem, or an integration of the joint direction-lattice distribution over the direction distribution induced by the sup-norm ball. We do not pursue this refinement here.

9 Finite-height linear independence

Define

$$L_2([v]) = \min\{|a|_2 : a \in \mathbb{Z}^{d+1} \setminus \{0\}, a \cdot v = 0\}.$$

Then

$$L_2([v]) = \lambda_1(\Lambda_v).$$

Corollary 9.1. For almost all $x \in \mathbb{P}^d(\mathbb{Q})$ with $H(x) \leq B$,

$$L_2(x) = B^{1/d+o(1)}.$$

Equivalently, for every $\varepsilon > 0$, almost every $x \in \mathbb{P}^d(\mathbb{Q})_{\leq B}$ satisfies no nontrivial rational hyperplane equation

$$a_0 X_0 + \cdots + a_d X_d = 0$$

with

$$|a|_2 \leq B^{1/d-\varepsilon}.$$

On the other hand, every $x \in \mathbb{P}^d(\mathbb{Q})_{\leq B}$ satisfies some such equation with

$$|a|_2 \ll_d B^{1/d}.$$

This is a finite-height form of linear algebraic independence from rational hyperplanes.

10 Relation with algorithmic statistics

In algorithmic statistics, a finite model for an object x is often represented by a finite set $A \ni x$. The two-part code length is

$$K(A) + \log |A|.$$

The structure function records the tradeoff between model complexity and residual index cost.

In the present setting, the model classes are arithmetic. A single rational hyperplane relation $a \cdot v = 0$ gives the finite model

$$A_{a,R} = \{v \in V_N(R) : a \cdot v = 0\}.$$

More globally, a cusp event

$$\lambda_1(\Lambda_v) \leq t|v|_2^{1/d}$$

gives the finite model

$$E_R(t) = \{v \in V_N(R) : \lambda_1(\Lambda_v) \leq t|v|_2^{1/d}\}.$$

The results above identify the natural volume of these models and hence their two-part coding consequences. A rational point with an unusually short relation or unusually short Bezout certificate lies in a small arithmetic model and is therefore Kolmogorov-deficient relative to the ambient height ball.

This is an arithmetic instance of the model-selection viewpoint of algorithmic statistics.

11 Further directions

11.1 Box-height limiting laws

The exponent statements are stable under replacing Euclidean height by projective height. A sharper box-height limiting law would require understanding the orthogonal-lattice distribution when primitive vectors are ordered by

$$|v|_\infty \leq B.$$

The joint direction-lattice equidistribution of Horesh and Karasik suggests that such a statement may be obtainable by integrating over the direction distribution for the cube.

11.2 Higher-degree relations

A natural next model class consists of bounded-degree hypersurfaces. Elementary counting shows that a fixed degree- D hypersurface in $\mathbb{P}^d$ contains $O_{d,D}(B^d)$ rational points of height at most B . Hence membership in a simple bounded-degree hypersurface forces at least $\log B$ bits of deficiency. A sharper problem is to study the minimal coefficient height of a degree- D polynomial relation satisfied by a typical rational point.

11.3 Inverse theorems

The forward implication is:

$$\text{simple arithmetic model} \Rightarrow \text{Kolmogorov deficiency}.$$

The inverse problem is much harder:

$$\text{persistent deficiency} \Rightarrow \text{membership in a simple arithmetic model}.$$

This cannot hold for unrestricted programs, but may hold relative to natural model classes such as rational hyperplanes, bounded-degree hypersurfaces, thin maps, or special subvarieties.

11.4 Frobenius certificates

For positive primitive vectors, Frobenius numbers also have limiting distributions governed by random lattices. This suggests extending the present relation-certificate framework from homogeneous relations and Bezout certificates to semigroup certificates.

References

- [HK] Y. Horesh and N. Karasik, *Equidistribution of primitive vectors, orthogonal lattices, and the shortest solution to the gcd equation*, arXiv:1903.01560.
- [LiV] M. Li and P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer.
- [Marklof] J. Marklof, *The asymptotic distribution of Frobenius numbers*, *Inventiones Mathematicae* 181 (2010), 179-207.
- [Rogers] C. A. Rogers, *Mean values over the space of lattices*, *Acta Mathematica* 94 (1955), 249-287.
- [Schmidt] W. M. Schmidt, *Diophantine Approximation*, *Lecture Notes in Mathematics* 785, Springer.
- [Siegel] C. L. Siegel, *A mean value theorem in geometry of numbers*, *Annals of Mathematics* 46 (1945), 340-347.
- [VV] N. Vereshchagin and P. Vitanyi, *Kolmogorov's structure functions and model selection*, *IEEE Transactions on Information Theory* 50 (2004), no. 12, 3265-3290.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, ChatGPT, by OpenAI, was used to assist with mathematical drafting, formalization, review, and editing. This work is shared as a preliminary AI-assisted mathematical note. The mathematical content may have been only partially reviewed and may contain errors; it should not be treated as peer-reviewed or as a fully verified manuscript.