ON SPARSE GEOMETRY OF NUMBERS

LENNY FUKSHANSKY, PAVEL GUERZHOY, AND STEFAN KÜHNLEIN

Abstract. Let L be a lattice of full rank in n-dimensional real space. A vector in L is called i-sparse if it has no more than i nonzero coordinates. We define the i-th successive sparsity level of L, $s_i(L)$, to be the minimal s so that L has s linearly independent i-sparse vectors, then $s_i(L) \leq n$ for each $1 \leq i \leq n$. We investigate sufficient conditions for $s_i(L)$ to be smaller than n and obtain explicit bounds on the sup-norms of the corresponding linearly independent sparse vectors in L. These results can be viewed as a partial sparse analogue of Minkowski's successive minima theorem. We then use this result to study virtually rectangular lattices, establishing conditions for the lattice to be virtually rectangular and determining the index of a rectangular sublattice. We further investigate the 2-dimensional situation, showing that virtually rectangular lattices in the plane correspond to elliptic curves isogenous to those with real j-invariant. We also identify planar virtually rectangular lattices in terms of a natural rationality condition of the geodesics on the modular curve carrying the corresponding points.

1. Introduction

Let $n \geq 2$ be an integer. For each $\boldsymbol{x} \in \mathbb{R}^n$, we write

$$\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}, \ |\boldsymbol{x}| = \max_{1 \leq i \leq n} |x_i|$$

for the usual Euclidean norm and sup-norm on $\mathbb{R}^n$, respectively. We also define the 0-*norm* on $\mathbb{R}^n$:

$$\|\boldsymbol{x}\|_0 := \sum_{i=1}^{n} x_i^0,$$

where we use the convention that $0^0 = 0$. The 0-norm counts the number of nonzero coordinates of a vector, which we refer to as the *sparsity level* of this vector; if sparsity level of some vector is no larger than m, we say that this vector is *m-sparse*. Sparsity has been actively investigated in the context of compressed sensing, which is a signal recovery paradigm based on the idea that most signals are sparse and can therefore be reconstructed from a small number of linear measurements [6]. More recently, the sparsity phenomenon has also been studied in discrete mathematics and discrete geometry, in particular in the context of lattices [7], [2], [1]. In this paper, we want to take a first stab at a systematic approach to what we see as a "sparse analogue" of the classical geometry of numbers.

2010 *Mathematics Subject Classification.* Primary: 11H06, 52C07, 11G05.

Key words and phrases. lattices, sparse vectors, virtually rectangular lattices, Siegel's lemma, elliptic curve, j-invariant, isogeny, modular curve, geodesics.

Fukshansky was partially supported by the Simons Foundation grant #519058.

Let $A = (a_{ij}) \in \mathrm{GL}_n(\mathbb{R})$, and define

$$|A| := \max_{1 \le i,j \le n} |a_{ij}|.$$

Let $L = A\mathbb{Z}^n \subset \mathbb{R}^n$, then L is a lattice of full rank with basis matrix A. The minimal norm of L is defined as

$$|L| := \min\{\|\boldsymbol{x}\| : \boldsymbol{x} \in L \setminus \{\mathbf{0}\}\}.$$

Previous research has focused on sparsity of integer representations of lattice vectors, i.e. on representing a vector $\boldsymbol{x} \in L$ as $\boldsymbol{x} = A\boldsymbol{y}$ with $\boldsymbol{y} \in \mathbb{Z}^n$ being as sparse as possible. In this paper, we will focus on the sparsity of the lattice vectors themselves. Specifically, we define the *successive sparsity levels* $s_1, \dots, s_n$ of the lattice L to be

$$\begin{aligned} s_i(L) \quad := \quad & \min\big\{s : \exists\ i \text{ linearly independent vectors } \boldsymbol{x}_1, \dots, \boldsymbol{x}_i \in L \\ & \text{with } \|\boldsymbol{x}_1\|_0, \dots, \|\boldsymbol{x}_i\|_0 \le s\big\}. \end{aligned}$$

Then $1 \le s_1 \le \cdots \le s_n \le n$. Given a lattice L, what can be said about its successive sparsity levels? Further, assuming we know that some $s_\ell \le k$, can we find ℓ such k-sparse vectors in L? To answer these questions, we need some more notation.

For every nonzero vector $\boldsymbol{x} \in L$ define

$$d(\boldsymbol{x}) := \dim_{\mathbb{Q}} \mathrm{span}_{\mathbb{Q}}\{x_1, \dots, x_n\}$$

to be its *rational dimension*. If A is an $n \times n$ real matrix with row vectors $\boldsymbol{a}_i$ for $1 \le i \le n$, then for each subset $I \subseteq [n] := \{1, \dots, n\}$ we define $d_I(A) := \sum_{i \in I} d(\boldsymbol{a}_i)$. We write $d(A)$ for $d_{[n]}(A)$, and define $d(L) = d(A)$, where A is any basis matrix for L. Indeed, this definition does not depend on the choice of a basis matrix: if A and B are two basis matrices for L, then $B = AU$ for some $U \in \mathrm{GL}_n(\mathbb{Z})$, and so each row vector $\boldsymbol{b}_i$ of B is of the form $\boldsymbol{b}_i = \boldsymbol{a}_i U$ for the corresponding row vector $\boldsymbol{a}_i$ of A, which implies that $d(\boldsymbol{b}_i) = d(\boldsymbol{a}_i)$. More generally, $d(A) = d(AU)$ holds for any $U \in \mathrm{GL}_n(\mathbb{Q})$. Notice that $d(L) \ge n$. We refer to $d(L)$ as the rational dimension of L: the smaller $d(L)$ is the "closer" L is to being *rational*, meaning $L \subset \mathbb{Q}^n$. Recall that L is *integral* if $\|\boldsymbol{x}\|^2 \in \mathbb{Z}$ for any $\boldsymbol{x} \in L$, and L is *arithmetic* if it is a scalar multiple of an integral lattice, so rational lattices are arithmetic. Certainly $d(L) = n$ for all rational lattices, but there also exist non-rational arithmetic lattices for which $d(L) = n$, for instance

$$L = \begin{pmatrix} \sqrt{2} & 2\sqrt{2} \\ \sqrt{2} & 3\sqrt{2} \end{pmatrix} \mathbb{Z}^2$$

is one such example. On the other hand, there exist arithmetic lattices with rational dimension $> n$, for instance

$$\begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & -1 \end{pmatrix} \mathbb{Z}^2, \ \begin{pmatrix} 1 & 1 & 1 \\ 1 & \sqrt{2} & 0 \\ \sqrt{2} & -1 & 0 \end{pmatrix} \mathbb{Z}^3,$$

and similar examples can be constructed in every dimension. Non-arithmetic lattices can also have rational dimension $> n$, for instance the planar lattice

$$\Lambda_1 = \begin{pmatrix} 1 & \sqrt{3} \\ 0 & 1 \end{pmatrix} \mathbb{Z}^2 \tag{1}$$

with $d(\Lambda_1) = 3$, as well as $= n$, for instance the planar lattice

$$\Lambda_2 = \begin{pmatrix} \pi & 2\pi \\ 2 & 1 \end{pmatrix} \mathbb{Z}^2$$

with $d(\Lambda_2) = 2$.

We will also define two "measures of irrationality" of vectors in L and of L itself. First, if $\boldsymbol{x} \in L$ has $d(\boldsymbol{x}) = k$, then it can be written as

$$\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i, \tag{2}$$

where $\boldsymbol{f}_1, \dots, \boldsymbol{f}_k$ are integer vectors with relatively prime coordinates. Notice that this decomposition is unique only if $k = 1$, for example

$$\begin{aligned} \left(1, \sqrt{2}, -2\sqrt{3}\right) &= 1 \cdot (1,0,0) + \sqrt{2} \cdot (0,1,0) - 2\sqrt{3} \cdot (0,0,1) \\ &= 1 \cdot (1,0,0) + \left(\frac{\sqrt{2}}{2} - \sqrt{3}\right) \cdot (0,1,1) + \left(\frac{\sqrt{2}}{2} + \sqrt{3}\right) \cdot (0,1,-1)\,. \end{aligned}$$

Then define

$$\nu(\boldsymbol{x}) := \begin{cases} |\alpha_1| & \text{if } k = 1, \\ 0 & \text{if } k > 1. \end{cases}$$

For our basis matrix A, we define

$$\nu(A) := \prod_{i=1}^{n} \nu(\boldsymbol{a}_i),$$

and for the lattice $L = A\mathbb{Z}^n$, we let $\nu(L) = \nu(A)$. This definition does not depend on the choice of the basis matrix A. Clearly, there are many lattices for which $\nu(L) = 0$, for example $\nu(\Lambda_1) = 0$, where Λ_1 is as in (1). In fact, it is not difficult to show that $\nu(L) > 0$ if and only if $d(L) = n$ (see Lemma 3.1 below).

Second, let $\langle L \rangle$ be the additive abelian group generated by the entries of vectors of L, and suppose that $\langle L \rangle$ has rank $k \geq 1$. Fix a basis $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)$ for $\langle L \rangle$ over $\mathbb{Z}$, then every $\boldsymbol{x} \in L$ has a representation of the form (2) with vectors $\boldsymbol{f}_1, \dots, \boldsymbol{f}_k \in \mathbb{Z}^n$. Define a map $\Phi_{\boldsymbol{\alpha}} : L \to \mathbb{R}^{nk}$ by

$$\Phi_{\boldsymbol{\alpha}} \left(\sum_{i=1}^{k} \alpha_i \boldsymbol{f}_i \right) = \begin{pmatrix} \boldsymbol{f}_1 \\ \vdots \\ \boldsymbol{f}_k \end{pmatrix}. \tag{3}$$

The map $\Phi_{\boldsymbol{\alpha}}$ is additive, and hence extends to a map $\mathbb{R} \otimes L = \mathbb{R}^n \to \mathbb{R}^{nk}$. We can then pull back the sup-norm $|\cdot|$ on $\mathbb{R}^{nk}$ to $\mathbb{R}^n$ under $\Phi_{\boldsymbol{\alpha}}$ by defining $|\boldsymbol{x}|_{\Phi_{\boldsymbol{\alpha}}} := |\Phi_{\boldsymbol{\alpha}}(\boldsymbol{x})|$, this way obtaining a norm $|\cdot|_{\Phi_{\boldsymbol{\alpha}}}$ on $\mathbb{R}^n$, which can then be compared to the sup-norm on $\mathbb{R}^n$. Specifically, we can define

$$\mu(\boldsymbol{\alpha}) := \sup \left\{ \frac{|\boldsymbol{x}|_{\Phi_{\boldsymbol{\alpha}}}}{|\boldsymbol{x}|} : \boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\} \right\}. \tag{4}$$

We can now state our first result.

Theorem 1.1. *Let $A \in \mathrm{GL}_n(\mathbb{R})$ and let $L = A\mathbb{Z}^n$. Fix a basis $\boldsymbol{\alpha}$ for $\langle L \rangle$ as above and let $\mu(\boldsymbol{\alpha})$ be as given in (4). Let $1 \leq k < n$ and suppose that there exists a subset $I \subset [n]$ of $n - k$ distinct indices such that $d_I(A) < n$. Let $\ell = n - d_I(A)$.*

Then $s_\ell(L) \leq k$, and there exist ℓ linearly independent vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_\ell \in L$ with $\|\boldsymbol{x}_i\|_0 \leq k$ and

$$\prod_{i=1}^{\ell} |\boldsymbol{x}_i| \leq n^{n-d_I(A)/2} |A|^n \mu(\boldsymbol{\alpha})^{d_I(A)}. \tag{5}$$

This theorem can be viewed as a "sparse" partial analogue of Minkowski's successive minima theorem. Indeed, if we know that $s_\ell(L) \leq k$, we can define the *k-sparse successive minima* $\lambda_1(L,k) \leq \cdots \leq \lambda_\ell(L,k)$ with respect to sup-norm to be

$$\lambda_i(L,k) := \min\left\{t \in \mathbb{R}_{>0} : \exists \text{ lin. ind. } \boldsymbol{x}_1, \dots, \boldsymbol{x}_i \in L \text{ with } \|\boldsymbol{x}_j\|_0 \leq k, |\boldsymbol{x}_j| \leq t\right\},$$

so the usual successive minima are $\lambda_i(L) := \lambda_i(L,n)$. Then (5) is an upper bound on the product of these k-sparse successive minima. We prove Theorem 1.1 in Section 2. Our main tool here is the celebrated Siegel's lemma, which is known to be sharp with respect to the exponent (we state it below as Theorem 2.3). We comment on the quality of the bound (5) at the end of Section 2. Unfortunately, the upper bound of (5) depends on the choice of the basis for L and for $\langle L \rangle$. We can alleviate this dependence for lattices L with rational dimension $d(L) = n$. We need some more notation.

Let us say that a lattice is *rectangular* if it has an orthogonal basis. Following [11], we will say that a lattice is *virtually rectangular* if it contains a rectangular sublattice of finite index. Two lattices $L, L' \subset \mathbb{R}^n$ are called *isometric* if there exists a real orthogonal matrix U such that $L' = UL$; on the other hand, L, L' are called *similar* if there exists a real orthogonal matrix U and a positive real number β such that $L' = \beta UL$. In other words, similarity as a linear map is a composition of an isometry and a dilation. It is easy to notice that the virtually rectangular property is preserved under isometry (and under similarity), however the sparsity levels, rational dimension and the irrationality measure ν are not necessarily preserved under isometry (they are preserved under dilation). In Section 3 we give the following characterization of virtually rectangular lattices, using the invariants we have just introduced.

Theorem 1.2. *Let $L \subset \mathbb{R}^n$ be a lattice of full rank. The following three statements are equivalent:*

1. $d(L) = n$,
2. $\nu(L) > 0$,
3. $s_1(L) = \cdots = s_n(L) = 1$.

Further, a full-rank lattice $L' \subset \mathbb{R}^n$ is virtually rectangular if and only if it is isometric to some lattice L satisfying the three equivalent conditions above.

In Section 3 we also prove the following effective result and show it to be optimal (Example 3.1).

Theorem 1.3. *Let $A \in \mathrm{GL}_n(\mathbb{R})$ be such that the lattice $L = A\mathbb{Z}^n$ satisfies the equivalent conditions of Theorem 1.2. Then L contains a rectangular sublattice M with a basis of 1-sparse vectors so that*

$$[L : M] = \left(\frac{\det(L)}{\nu(L)}\right)^{n-1}. \tag{6}$$

More generally, if $L' \subset \mathbb{R}^n$ be a virtually rectangular lattice, then there exists a rectangular sublattice M' of L' such that $[L' : M'] = \left(\frac{\det(L)}{\nu(L)}\right)^{n-1}$, where L is a lattice isometric to L' which satisfies the equivalent conditions of Theorem 1.2.

In the 2-dimensional case our results imply a certain property of elliptic curves over $\mathbb{C}$. An elliptic curve E can be realized as a complex torus $\mathbb{C}/\Lambda$ for a planar lattice Λ, called the *period lattice* of this curve. Given two elliptic curves, E and E' a morphism $\phi : E \to E'$ between them such that $\phi(0) = 0$ and $\phi(E) \neq \{0\}$ is called an *isogeny*. It is a remarkable fact that an isogeny is always surjective and has a finite kernel, the order of which is called the *degree* of the isogeny, denoted $\delta(E/E')$. If an isogeny $E \to E'$ exists, then there also exists the dual isogeny $E' \to E$ of the same degree such that their composition is simply the multiplication-by-$\delta(E/E')$ map, and hence the curves are called *isogenous*: this is an equivalence relation. An injective isogeny is called an *isomorphism*, and the set of isomorphism classes of elliptic curves over $\mathbb{C}$ is parameterized by

$$\mathcal{D} := \{\tau = a+bi \in \mathbb{C} : -1/2 < a \leq 1/2, b \geq 0, |\tau| \geq 1\} \backslash \{e^{i\theta} : \pi/2 < \theta < 2\pi/3\} \tag{7}$$

in the following way. For each $\tau = a + bi \in \mathcal{D}$ we can define a lattice

$$\Gamma_\tau = \mathbb{Z} + \mathbb{Z}\tau, \tag{8}$$

which can be thought of as $\begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix} \mathbb{Z}^2$ in $\mathbb{R}^2$. Then every elliptic curve is isomorphic to an elliptic curve E_τ with period lattice Γ_τ for some $\tau \in \mathcal{D}$. In fact, it is more natural to identify the set of isomorphic classes of elliptic curves with the quotient space of the upper half-plane under the action of $\mathrm{SL}_2(\mathbb{Z})$ by linear fractional transformations, where $\mathcal{D}$ is a fundamental domain for this action. There is a unique bijective holomorphic map $j : \mathcal{D} \to \mathbb{C}$ taking $e^{2\pi i/3}$ to 0 and i to 1728, called the *Klein j-function*, which is modular and gives the *j-invariant* $j(\tau)$ for each isomorphism class E_τ of elliptic curves. In [8] the relevant properties of the j-invariant are outlined, and in particular it is noted that for $\tau \in \mathcal{D}$, $j(\tau) \in \mathbb{R}$ if and only if τ belongs to the set

$$\left\{1/2 + it : t \in \mathbb{R}, t \geq \sqrt{3}/2\right\} \cup \left\{e^{i\theta} : \theta \in [\pi/3, \pi/2]\right\} \cup \left\{it : t \in \mathbb{R}, t \geq 1\right\}, \tag{9}$$

and j maps the first of these three subsets bijectively onto the interval $(-\infty, 0]$, the second onto $[0, 1]$, and the third onto $[1, \infty)$ (see also Proposition on p. 160 of [10] for an earlier appearance of this observation). With this notation, we can state the following result which we prove in Section 4.

Theorem 1.4. *Let $\tau = a + bi \in \mathcal{D}$ and let E_τ be the corresponding elliptic curve with the period lattice Γ_τ as above. The following statements are equivalent:*

1. *Either $a \in \mathbb{Q}$ or there exists some $t \in \mathbb{R}$ such that $a - bt, a + b/t \in \mathbb{Q}$,*
2. *Γ_τ is virtually rectangular,*
3. *E_τ is isogenous to an elliptic curve E' with real j-invariant ≥ 1,*
4. *E_τ is isogenous to an elliptic curve E' with real j-invariant in $[0, 1]$.*

If these equivalent conditions hold with $a \in \mathbb{Q}$, then there exists such an isogeny $E' \to E_\tau$ with $\delta(E'/E_\tau) =$ the denominator of a. If the conditions hold with $a \notin \mathbb{Q}$ and t is any real number satisfying (1), then there exists such an isogeny $E' \to E_\tau$

with

$$\delta(E'/E_\tau) = \frac{|b|vw(t^2+1)}{|t|}, \tag{10}$$

where $v, w > 0$ are denominators of the rational numbers $a - bt$ and $a + b/t$, respectively.

Our proof of this theorem uses Theorem 1.2. In particular, condition (1) of Theorem 1.4 is equivalent to condition (1) of Theorem 1.2 in this 2-dimensional situation. Further, (10) is just a reformulation of (6) in this case, since $\delta(E'/E_\tau)$ is precisely the index of the rectangular period lattice of E' as a sublattice in the virtually rectangular period lattice Γ_τ of E_τ. In the case when $a = p/q \in \mathbb{Q}$, $d(\Gamma_\tau) = 2$.

We will refer to elliptic curves satisfying the conditions of Theorem 1.4 as *virtually rectangular*. The class of their period lattices includes all Γ_τ so that $j(\tau) \in \mathbb{R}$ (see (9)). Further, this class includes all arithmetic planar lattices (see Lemma 2.5 of [11]), which are Γ_τ for $\tau \in \mathcal{D}$ being a quadratic irrationality (see [8]): these are $\tau = a + bi$ with $a, b^2 \in \mathbb{Q}$, which correspond precisely to elliptic curves with complex multiplication (CM). We will discuss this situation in more details in Section 4, in particular proving that CM elliptic curves are the only ones whose period lattice contains non-parallel rectangular sublattices (Proposition 4.2 and Corollary 4.3): in the CM case, there are infinitely many t satisfying condition (1) of Theorem 1.4 (each corresponding to a different rectangular sublattice), whereas for all other virtually rectangular elliptic curves such t is essentially unique. Finally, in Section 5 we will show that virtually rectangular lattices in the plane have intrinsic geometric meaning in terms of the corresponding points on the modular curve: they correspond precisely to the points that lie on geodesics closed at infinity (Theorem 5.1).

2. Successive sparsity levels

In this section we prove Theorem 1.1. We start with a lemma on successive sparsity levels.

Lemma 2.1. *Let $A \in \mathrm{GL}_n(\mathbb{R})$ and let $L = A\mathbb{Z}^n$. Let $1 \le k < n$ and suppose that there exists a subset $I \subset \{1, \dots, n\}$ of $n - k$ distinct indices such that $d_I(A) < n$. Let $\ell = n - d_I(A)$. Then $s_\ell(L) \le k$.*

Proof. Let $I = \{i_1, \dots, i_{n-k}\}$ for some $1 \le i_1 < i_2 < \cdots < i_{n-k} \le n$, and let us write $d_j := d(\boldsymbol{a}_{i_j})$ for each $1 \le j \le n-k$. Let A_I be the $(n-k) \times n$ submatrix of A consisting of the rows indexed by I. We want to show that there exists a nonzero vector $\boldsymbol{x} \in \mathbb{Z}^n$ such that $A_I \boldsymbol{x} = \boldsymbol{0}$. For each $1 \le j \le n-k$, let

$$V_j = \left\{ \boldsymbol{x} \in \mathbb{Q}^n : \boldsymbol{a}_{i_j} \cdot \boldsymbol{x} = 0 \right\},$$

then $\dim_{\mathbb{Q}} V_j = n - d_j$. Further, let us prove that

$$\dim_{\mathbb{Q}} \left(\bigcap_{j=1}^{n-k} V_j \right) \ge \ell = n - d_I(A).$$

We argue by induction on $n - k \ge 1$. If $n - k = 1$, then $I = \{i_1\}$ and so $d_I(A) = d_1$, in which case

$$\dim_{\mathbb{Q}} V_1 = n - d_1 = \ell.$$

Now assume the result for all $1 \le n-k < m \le n-1$, and let us prove it for $n-k=m$. Let

$$I' = \{i_1, \dots, i_{m-1}\}, \text{ so } I = I' \cup \{i_m\},$$

then $d_I(A) = d(I') + d_m$. Let $V' = \bigcap_{j=1}^{m-1} V_j$ and $V = V' \cap V_m$. By induction hypothesis,

$$\dim_{\mathbb{Q}} V' \ge n - d(I').$$

Since $d(I') < d_I(A) < n$, this implies that $\dim_{\mathbb{Q}} V' > 0$, and so $V' \ne \{\mathbf{0}\}$. Now, by a well-known identity in linear algebra,

$$\begin{aligned} \dim_{\mathbb{Q}} V &= \dim_{\mathbb{Q}} V' + \dim_{\mathbb{Q}} V_m - \dim_{\mathbb{Q}} \operatorname{span}_{\mathbb{Q}}\{V', V_m\} \\ &\ge (n - d(I')) + (n - d_m) - \dim_{\mathbb{Q}} \operatorname{span}_{\mathbb{Q}}\{V', V_m\} \\ &\ge n - d_I(A) = \ell, \end{aligned}$$

since $\operatorname{span}_{\mathbb{Q}}\{V', V_m\} \subseteq \mathbb{Q}^n$, and so $\dim_{\mathbb{Q}} \operatorname{span}_{\mathbb{Q}}\{V', V_m\} \le n$.

This implies that $\dim_{\mathbb{Q}} \left(\bigcap_{j=1}^{n-k} V_j\right) = \ell > 0$, and so there exist ℓ nonzero linearly independent vectors $\boldsymbol{y}_1, \dots, \boldsymbol{y}_\ell \in \bigcap_{j=1}^{n-k} V_j$. These vectors are in $\mathbb{Q}^n$ and satisfy the equation $A_I \boldsymbol{y}_i = \mathbf{0}$. Multiplying $\boldsymbol{y}_1, \dots, \boldsymbol{y}_\ell$ by the least common denominator of their coordinates, we obtain linearly independent vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_\ell \in \mathbb{Z}^n$ such that $A_I \boldsymbol{x}_i = \mathbf{0}$. This means that the vectors $A\boldsymbol{x}_1, \dots, A\boldsymbol{x}_\ell \in L$ have at least $n-k$ coordinates equal to 0. Since $\boldsymbol{x}_1, \dots, \boldsymbol{x}_\ell$ are linearly independent and A is a nonsingular matrix, we must have $A\boldsymbol{x}_1, \dots, A\boldsymbol{x}_\ell$ linearly independent, and so $s_\ell(L) \le k$. □

Remark 2.1. Notice that

$$d_I(A) \ge \dim_{\mathbb{Q}} A_I := \dim_{\mathbb{Q}}\{a_{i_1 1}, \dots, a_{i_k n}\}.$$

The converse of Lemma 2.1 is not true: if $s_1(L) = k$, there may not exist any $I \subset \{1, \dots, n\}$ of cardinality $n-k$ so that $d_I(A) < n$. Indeed, consider the example

$$A = \begin{pmatrix} 1 & \sqrt{3} & 2\sqrt{3} \\ \sqrt{5} & \sqrt{3} & 2\sqrt{3} \\ \sqrt{2} & \sqrt{3} & \sqrt{5} \end{pmatrix} \tag{11}$$

and let $L = A\mathbb{Z}^3$. Then $s_1(L) = 1$, since

$$A \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2\sqrt{3} - \sqrt{5} \end{pmatrix} \in L.$$

On the other hand, $d(\boldsymbol{a}_1) = d(\boldsymbol{a}_2) = 2$ and $d(\boldsymbol{a}_3) = 3$. Thus for any I of cardinality $3 - 1 = 2$, $d_I(A) \ge 4 > n = 3$. Further, even $\dim_{\mathbb{Q}} A_I$ here is at least 3. On the other hand, notice for comparison purposes that if a lattice $L = A\mathbb{Z}^n$ is virtually rectangular, then $\dim_{\mathbb{Q}}(A^\top A) \le n$ (see [11]).

For each row vector $\boldsymbol{a}_i$ of A let $d_i = d(\boldsymbol{a}_i)$. Then there exist $\mathbb{Q}$-linearly independent real numbers $\alpha_{i1}, \dots, \alpha_{id_i}$ such that

$$\boldsymbol{a}_i = \sum_{j=1}^{d_i} \alpha_{ij} \boldsymbol{f}_{ij}, \tag{12}$$

where $\boldsymbol{f}_{ij}$ are integer vectors with relatively prime coefficients for all $1 \le i \le n$, $1 \le j \le d_i$. Let $d = \sum_{i=1}^n d_i$ and let $F(A)$ be the $d \times n$ matrix with rows $\boldsymbol{f}_{ij}$.

Example 2.1. For instance, in case of the matrix A in (11), we have $d_1 = 2$, $d_2 = 2$, $d_3 = 3$, and

$$\boldsymbol{a}_1 = \alpha_{11}\boldsymbol{f}_{11} + \alpha_{12}\boldsymbol{f}_{12},\ \boldsymbol{a}_2 = \alpha_{21}\boldsymbol{f}_{21} + \alpha_{22}\boldsymbol{f}_{22},\ \boldsymbol{a}_3 = \alpha_{31}\boldsymbol{f}_{31} + \alpha_{32}\boldsymbol{f}_{32} + \alpha_{33}\boldsymbol{f}_{33},$$

where

$$\alpha_{11} = 1,\ \alpha_{12} = \alpha_{22} = \alpha_{32} = \sqrt{3},\ \alpha_{21} = \alpha_{33} = \sqrt{5},\ \alpha_{31} = \sqrt{2},$$

and

$$\boldsymbol{f}_{11} = \boldsymbol{f}_{21} = \boldsymbol{f}_{31} = (1,0,0),\ \boldsymbol{f}_{12} = \boldsymbol{f}_{21} = (0,1,2),\ \boldsymbol{f}_{32} = (0,1,0),\ \boldsymbol{f}_{33} = (0,0,1).$$

Therefore $d = 2 + 2 + 3 = 7$ in this example, and the 7×3 matrix $F(A)$ is

$$F(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Define

$$|F(A)| := \max_{1\le i\le n} \max_{1\le j\le d_i} |\boldsymbol{f}_{ij}|,$$

where $|\boldsymbol{f}_{ij}|$ is the sup-norm of the vector $\boldsymbol{f}_{ij}$.

Lemma 2.2. *For a fixed choice of $\boldsymbol{\alpha} = (\alpha_{ij} : 1 \le i \le n, 1 \le j \le d_i)$ as above, we have*

$$|F(A)| \le \mu(\boldsymbol{\alpha})|A|,$$

where $\mu(\boldsymbol{\alpha})$ is as in (4).

Proof. Notice that the rank of the additive group $\langle L\rangle$ is equal to d, and $\boldsymbol{\alpha}$ is a basis for it. Using the notation of Section 1, specifically (3), notice that for each $1 \le i \le n$,

$$\frac{|\boldsymbol{a}_i|_{\Phi_{\boldsymbol{\alpha}}}}{|\boldsymbol{a}_i|} \le \mu(\boldsymbol{\alpha}).$$

On the other hand, $|\boldsymbol{a}_i|_{\Phi_{\boldsymbol{\alpha}}} = |\Phi_{\boldsymbol{\alpha}}(\boldsymbol{a}_i)| = \max_{1\le j\le d_i} |\boldsymbol{f}_{ij}|$, and hence

$$|F(A)| = \max_{1\le i\le n} \max_{1\le j\le d_i} |\boldsymbol{f}_{ij}| \le \mu(\boldsymbol{\alpha}) \max_{1\le i\le n} |\boldsymbol{a}_i| = \mu(\boldsymbol{\alpha})|A|.$$

□

Our next result relies heavily on the use of Siegel's lemma (Theorem 2 of [4]) and its adaptation to sup-norm using Fisher's inequality (equation (1.8) of [3]). We state it here for the reader's convenience.

Theorem 2.3 (Siegel's lemma with matrix sup-norm)**.** *Let B be an $m \times n$ integer matrix of rank $m < n$. Then there exist $n - m$ linearly independent vectors $\boldsymbol{z}_1, \dots, \boldsymbol{z}_{n-m} \in \mathbb{Z}^n$ such that $B\boldsymbol{z}_i = \boldsymbol{0}$ for every $1 \le i \le n - m$ and*

$$\prod_{i=1}^{n-m} |\boldsymbol{z}_i| \le \left(\sqrt{n}|B|\right)^m.$$

The exponent m in this upper bound cannot in general be improved.

Lemma 2.4. *Let $A \in \mathrm{GL}_n(\mathbb{R})$, $L = A\mathbb{Z}^n$ and let $\boldsymbol{\alpha}$ be a fixed basis for $\langle L \rangle$. Let $1 \leq k < n$ and suppose that there exists a subset $I \subset \{1, \dots, n\}$ of $n-k$ distinct indices*

$$1 \leq i_1 < i_2 < \cdots < i_{n-k}$$

such that $d_I(A) := \sum_{j=1}^{n-k} d_{i_j} < n$. Let $\ell = n - d_I(A)$. Then there exist ℓ linearly independent vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_\ell \in L$ with $\|\boldsymbol{x}_i\|_0 \leq k$ and

$$\prod_{i=1}^{\ell} |\boldsymbol{x}_i| \leq n^{n-d_I(A)/2} |A|^n \mu(\boldsymbol{\alpha})^{d_I(A)}.$$

Proof. Let A_I be the $(n-k) \times n$ submatrix of A consisting of the rows indexed by I. Let $F(A)_I$ be the $d_I(A) \times n$ submatrix of $F(A)$ consisting of rows with $\boldsymbol{f}_{i_l j}$ for $i_l \in I$ and $1 \leq j \leq d_{i_l}$. Notice that $A_I \boldsymbol{y} = \boldsymbol{0}$ for some $\boldsymbol{y} \in \mathbb{Z}^n$ if and only if $F(A)_I \boldsymbol{y} = 0$. Using the notation in the proof of Lemma 2.1, we have

$$V = \left(\bigcap_{j=1}^{n-k} V_j \right) = \{ \boldsymbol{y} \in \mathbb{Q}^n : A_I \boldsymbol{y} = \boldsymbol{0} \} = \{ \boldsymbol{y} \in \mathbb{Q}^n : F(A)_I \boldsymbol{y} = \boldsymbol{0} \},$$

which is an ℓ-dimensional subspace of $\mathbb{Q}^n$. By Theorem 2.3, there exist ℓ linearly independent vectors $\boldsymbol{y}_1, \dots, \boldsymbol{y}_\ell \in V \cap \mathbb{Z}^n$ such that

$$\prod_{j=1}^{\ell} |\boldsymbol{y}_j| \leq \left(\sqrt{n} |F(A)| \right)^{d_I(A)}. \tag{13}$$

For each $1 \leq j \leq \ell$, define $\boldsymbol{x}_j = A\boldsymbol{y}_j$. Since A is a nonsingular matrix, $\boldsymbol{x}_1, \dots, \boldsymbol{x}_\ell$ are nonzero linearly independent vectors in L which are at least k-sparse. Now notice that

$$|\boldsymbol{x}_j| = |A\boldsymbol{y}_j| \leq n|A||\boldsymbol{y}_j|, \tag{14}$$

and so

$$\prod_{i=1}^{\ell} |\boldsymbol{x}_i| \leq n^\ell |A|^\ell \prod_{j=1}^{\ell} |\boldsymbol{y}_j|.$$

Combining this observation with (13) and Lemma 2.2 completes the proof. □

Theorem 1.1 now follows by combining Lemmas 2.1 and 2.4. Notice that the exponent $d_I(A)$ in the upper bound of (13) is sharp due to the optimality of the bound in Theorem 2.3. On the other hand, dependence of the bound of (14) on $|A|$ also cannot be improved in general. This suggests that the dependence of the bound of Theorem 1.1 on $|A|$ and $\mu(\boldsymbol{\alpha})$ has correct order of magnitude.

3. Virtually rectangular lattices

In this section we focus on virtually rectangular lattices. We start by presenting the proof of Theorem 1.2, split into two parts.

Lemma 3.1. *Let $L \subset \mathbb{R}^n$ be a lattice of full rank. The following three statements are equivalent:*

1. *$d(L) = n$,*
2. *$\nu(L) > 0$,*
3. *$s_1(L) = \cdots = s_n(L) = 1$.*

Proof. Let $L = A\mathbb{Z}^n$, where A is a basis matrix with rows $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$. We will prove that (1) is equivalent to (2) and that (1) is equivalent to (3). First assume that $d(L) = n$, then $d(\boldsymbol{a}_i) = 1$ for each row $\boldsymbol{a}_i$ of the basis matrix A. This means that each $\boldsymbol{a}_i = \alpha_i \boldsymbol{z}_i$, where $\alpha_i \in \mathbb{R} \setminus \{0\}$ and $\boldsymbol{z}_i \in \mathbb{Z}^n$ is a vector with relatively prime coordinates, and so $\nu(\boldsymbol{a}_i) = |\alpha_i|$. Then

$$\nu(L) = \nu(A) = \prod_{i=1}^{n} |\alpha_i| > 0,$$

and so (1) implies (2). Further, this means that for any subset $I \subset [n]$ of cardinality $n-1$, the linear system $A_I \boldsymbol{y} = \boldsymbol{0}$ has a nontrivial integer solution. Since such sets I are of the form $I = [n] \setminus \{i\}$, $1 \le i \le n$, we can index corresponding integer solutions by $\boldsymbol{y}_i$. Then each vector $A\boldsymbol{y}_i$ has only the i-th coordinate nonzero, and hence all of such vectors are linearly independent (they are multiples of the standard basis vectors). Therefore $s_1(L) = \cdots = s_n(L) = 1$, and so (1) implies (3).

Next assume $\nu(L) > 0$, and let A be a basis matrix for L. Then

$$\nu(L) = \nu(A) = \prod_{i=1}^{n} \nu(\boldsymbol{a}_i) > 0,$$

which means that each $\boldsymbol{a}_i$ is of the form $\boldsymbol{a}_i = \alpha_i \boldsymbol{z}_i$ for some $\alpha_i \in \mathbb{R} \setminus \{0\}$ and $\boldsymbol{z}_i \in \mathbb{Z}^n$ a primitive vector. Hence $d(L) = n$, and so (2) implies (1).

Finally, suppose $s_1(L) = \cdots = s_n(L) = 1$. Then there exist linearly independent vectors $a_1\boldsymbol{e}_1, \dots, a_n\boldsymbol{e}_n \in L$, where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ are the standard basis vectors. Hence there exists $U \in \mathrm{GL}_n(\mathbb{Q})$ such that AU is a nonsingular diagonal matrix, which implies $d(A) = d(AU) = n$. Thus (3) implies (1). □

Lemma 3.2. *A lattice L is virtually rectangular if and only if it is isometric to some lattice L' with $s_1(L') = \cdots = s_n(L') = 1$.*

Proof. Suppose that L contains a rectangular sublattice M, and let B be an orthogonal basis matrix for M. Then there exists a real orthogonal matrix U such that UB is a diagonal matrix. Let $M' = UM = UB\mathbb{Z}^n$ be a sublattice of the lattice $L' = UL$. Since UB is diagonal, M' has a basis consisting of scalar multiples of the standard basis vectors, and thus all successive sparsity levels of L' are equal to 1.

Conversely, assume L is isometric to some L' with successive sparsity levels equal to 1, say $L = UL'$ for some orthogonal matrix U. Then L' contains n linearly independent vectors $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ with $\|\boldsymbol{x}_i\|_0 = 1$. These vectors must therefore be constant multiples of standard basis vectors. Let $M' = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$, then $M = UM'$ is a rectangular sublattice of L. □

Then Theorem 1.2 follows by combining Lemmas 3.1 and 3.2. Next we prove Theorem 1.3.

Proof of Theorem 1.3. Since $d(L) = n$, we must have $d(\boldsymbol{a}_i) = 1$ for each row vector $\boldsymbol{a}_i$ of A. Then there must exist nonzero real numbers $\alpha_1, \dots, \alpha_n$ and primitive integer row vectors $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ so that $\boldsymbol{a}_i = \alpha_i \boldsymbol{f}_i$ for each $1 \le i \le n$. Hence the matrix $F(A)$ with row vectors $\boldsymbol{f}_i$ is $n \times n$ and $A = \mathcal{A}F(A)$, where $\mathcal{A}$ is the diagonal matrix with diagonal entries $\alpha_1, \dots, \alpha_n$. Let $\mathrm{adj}(F(A))$ be the adjugate of $F(A)$, then $\mathrm{adj}(F(A))$ is an integer matrix in $\mathrm{GL}_n(\mathbb{Q})$, $\det(\mathrm{adj}(F(A))) = \det(F(A))^{n-1}$ and

$$F(A)\,\mathrm{adj}(F(A)) = \det(F(A))I_n,$$

where I_n is the $n \times n$ identity matrix. Then

$$B := A \operatorname{adj}(F(A)) = \mathcal{A}F(A) \operatorname{adj}(F(A)) = \det(F(A))\mathcal{A},$$

which is a diagonal matrix with diagonal entries $\alpha_1 \det(F(A)), \ldots, \alpha_n \det(F(A))$. This implies that

$$\det(B) = \det(F(A))^n \det(\mathcal{A}) = \det(F(A))^n \prod_{i=1}^{n} \alpha_i.$$

On the other hand, since $\operatorname{adj}(F(A))$ is an integer matrix in $\mathrm{GL}_n(\mathbb{Q})$, the lattice $M := B\mathbb{Z}^n$ is a full-rank sublattice of L, which is rectangular with a basis of 1-sparse vectors. Further,

$$\begin{aligned} [L : M] &= \frac{\det(M)}{\det(L)} = \frac{|\det(B)|}{|\det(A)|} = |\det(\operatorname{adj}(F(A))| \\ &= |\det(F(A))|^{n-1} = \left|\frac{\det(A)}{\det(\mathcal{A})}\right|^{n-1} = \left(\frac{\det(L)}{\nu(L)}\right)^{n-1}. \end{aligned}$$

Now, if $L' \subset \mathbb{R}^n$ is any virtually rectangular lattice, then it is isometric to some lattice L satisfying the equivalent conditions of Theorem 1.2. This L has a rectangular sublattice M as we just constructed satisfying (6). Then L' contains a rectangular sublattice M' isometric to M with $[L' : M'] = [L : M]$. □

Example 3.1. Theorem 1.3 is optimal, i.e. the lattice $L = A\mathbb{Z}^n$ as in the statement of the theorem may not contain a rectangular sublattice M with a basis of 1-sparse vectors and smaller index than given in (6). Indeed, this is easily seen to be the case when

$$A = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ 1 & 1 & \ldots & 1 & d \end{pmatrix}$$

for some integer $d > 1$. Here $\nu(L) = 1$, $\det(L) = d$, and the smallest-index rectangular sublattice with a basis of 1-sparse vectors is

$$M = \begin{pmatrix} d & 0 & \ldots & 0 & 0 \\ 0 & d & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & d & 0 \\ 0 & 0 & \ldots & 0 & d \end{pmatrix} \mathbb{Z}^n,$$

which has index d^{n-1} in L.

Remark 3.1. It may be instructive to separately consider the 2-dimensional case of Theorem 1.3, where the computation becomes completely elementary. Let $L' \subset \mathbb{R}^2$ be a virtually rectangular lattice and let L be a lattice isometric to L' which satisfies the equivalent conditions of Theorem 1.2. Then $L = A\mathbb{Z}^2$, where

$$A = \begin{pmatrix} \alpha_1 u_1 & \alpha_1 v_1 \\ \alpha_2 u_2 & \alpha_2 v_2 \end{pmatrix}$$

with u_1, u_2, v_1, v_2 relatively prime integers and α_1, α_2 real numbers. Then

$$\det(L) = \alpha_1 \alpha_2 |u_1 v_2 - u_2 v_1|, \ \nu(L) = |\alpha_1 \alpha_2|.$$

Further, it is easy to see that the orthogonal vectors

$$\boldsymbol{z}_1 = \begin{pmatrix} 0 \\ \alpha_2(u_1v_2 - u_2v_1) \end{pmatrix}, \ \boldsymbol{z}_2 = \begin{pmatrix} \alpha_1(u_2v_1 - u_1v_2) \\ 0 \end{pmatrix}$$

are in L, and so $M = \mathrm{span}_{\mathbb{Z}}\{\boldsymbol{z}_1, \boldsymbol{z}_2\}$ is a rectangular sublattice of L. Then

$$\det(M) = |\alpha_1\alpha_2(u_1v_2 - u_2v_1)(u_2v_1 - u_1v_2)| = \frac{\det(L)^2}{|\alpha_1\alpha_2|}.$$

Let M' be a sublattice of L' isometric to M in L, then

$$[L' : M'] = [L : M] = \frac{\det(M)}{\det(L)} = \frac{\det(L)}{\nu(L)}.$$

4. Isogenies of elliptic curves

In this section we prove Theorem 1.4 and discuss some of its consequences. To start with, we state a technical lemma that will be of use to us: it is a combination of Lemma 5.3 and Proposition 5.4 of [8] (see also Proposition on p. 160 of [10], as mentioned in Section 1).

Lemma 4.1. *Let $\mathcal{D}$ be as in* (7)*. For $\tau \in \mathcal{D}$, the value $j(\tau)$ is real if and only if τ belongs to the set described in* (9)*. Further, Γ_τ is WR if and only if $j(\tau)$ is real and belongs to the interval $[0, 1]$.*

Proof of Theorem 1.4. First suppose that $a = \frac{p}{q} \in \mathbb{Q}$, then Γ_τ contains orthogonal vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 0 \\ qb \end{pmatrix} = q \begin{pmatrix} a \\ b \end{pmatrix} - p \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

These two vectors span a rectangular sublattice of Γ_τ of determinant qb, i.e. of index q, which in particular implies that Γ_τ is virtually rectangular. If $a \notin \mathbb{Q}$, assume that there exists some $t \in \mathbb{R}$ such that $a - bt, a + b/t \in \mathbb{Q}$. Define the lattice

$$L_t := \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} 1 & a - bt \\ t & at + b \end{pmatrix} \mathbb{Z}^2,$$

then it is easy to see that $d(L_t) = 2$, and so it is virtually rectangular by Theorem 1.2. Let $\theta = \arctan t$, then $\cos\theta = \frac{1}{\sqrt{1+t^2}}$ and $\sin\theta = \frac{t}{\sqrt{1+t^2}}$, meaning that

$$U_t = \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} 1 & -t \\ t & 1 \end{pmatrix}$$

is an orthogonal matrix. Notice that $U_t\Gamma_\tau = L_t$, meaning that Γ_τ is isometric to L_t, hence it is also virtually rectangular. This shows that condition (1) implies (2).

Suppose now Γ_τ is virtually rectangular, then it contains a rectangular sublattice Γ'. Let E' be the elliptic curve (up to isomorphism) with period lattice Γ' (up to similarity). We can then assume that

$$\Gamma' = \begin{pmatrix} 1 & 0 \\ 0 & q \end{pmatrix} \mathbb{Z}^2, \tag{15}$$

which means that $E' = E_{\tau'}$ for $\tau' = iq$. Hence τ' is in the third component set of (9), and so $j(\tau') \geq 1$ (see Lemma 4.1 above). Now, since the period lattice of E' is a sublattice of the period lattice of E_τ, there must exist an isogeny $E' \to E_\tau$ induced by the projection $\mathbb{C}/\Gamma' \to \mathbb{C}/\Gamma_\tau$. This shows that condition (2) implies (3).

An isogeny $E' \to E$ exists if and only if the period lattice for E' is (up to similarity) a sublattice of the period lattice for E. A planar lattice is called *well-rounded (WR)* if it has two linearly independent shortest vectors with respect to Euclidean norm. Now, the period lattice is rectangular if and only if the corresponding j-invariant is real and ≥ 1 while the period lattice is WR if and only if the corresponding j-invariant is real and in the interval $[0, 1]$ (Lemma 4.1). Hence to prove that conditions (3) and (4) are equivalent it is sufficient to show that Γ_τ contains a rectangular sublattice if and only if it contains a WR sublattice. This is guaranteed by Lemma 2.1 of [11].

Next assume that E_τ is isogenous to some elliptic curve $E' = E_{\tau'}$ with real nonnegative j-invariant. Let $\Gamma' = \Gamma_{\tau'}$ be the period lattice of E', so $j(\tau') \in \mathbb{R}_{\geq 0}$ and Γ' is (up to similarity) a sublattice of Γ. If $j(\tau') \geq 1$, then Lemma 4.1 guarantees that $\tau' = iq$ for some $q \geq 1$, and so Γ' is of the form (15), which is rectangular. If, on the other hand, $0 \leq j(\tau') < 1$, then Lemma 4.1 implies that the lattice Γ' is WR. Now, Lemma 2.1 of [11] asserts that a lattice has WR sublattices if and only if it is virtually rectangular. Hence we conclude in any case that Γ_τ is virtually rectangular. Therefore Γ_τ is isometric to some lattice L with $d(L) = 2$, by Theorem 1.2. Let

$$U(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

for some angle θ be the corresponding isometry matrix, and let $\tau = a + bi$ so that Γ_τ is of the form (8). Then

$$L = U(\theta)\Gamma_\tau = \begin{pmatrix} \cos\theta & a\cos\theta - b\sin\theta \\ \sin\theta & a\sin\theta + b\cos\theta \end{pmatrix} \mathbb{Z}^2 = \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} 1 & a - bt \\ t & at + b \end{pmatrix} \mathbb{Z}^2,$$

where $t = \tan\theta$. Since $d(L) = 2$, we must have $a - bt \in \mathbb{Q}$ and $\frac{at+b}{t} = a + b/t \in \mathbb{Q}$. This shows that condition (3) implies (1).

Finally, assume that the equivalent conditions of Theorem 1.4 hold. If $a = \frac{p}{q} \in \mathbb{Q}$, then Γ_τ contains a rectangular sublattice Γ' of index q. Let E' be the elliptic curve (up to isomorphism) corresponding to Γ', then the degree of the isogeny $E' \to E_\tau$ is precisely this index q. If $a \notin \mathbb{Q}$, then equivalent conditions of Theorem 1.4 hold with some $t \in \mathbb{R}$. Then the period lattice Γ_τ of the curve E_τ is virtually rectangular and isometric to the lattice $L_t = A_t\mathbb{Z}^2$ with

$$A_t = \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} 1 & a - bt \\ t & at + b \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+t^2}} & \frac{1}{\sqrt{1+t^2}}(a - bt) \\ \frac{t}{\sqrt{1+t^2}} & \frac{t}{\sqrt{1+t^2}}\left(a + \frac{b}{t}\right) \end{pmatrix},$$

and $d(L'_t) = 2$. Since $a - bt, a + b/t \in \mathbb{Q}$, we can write

$$a - bt = \frac{u}{v},\ a + \frac{b}{t} = \frac{q}{w}$$

with $u, v, q, w \in \mathbb{Z}$ and $v, w > 0$. Then, repeating the argument of Remark 3.1 for this specific situation,

$$u - (a - bt)v = 0,\ q - \left(a + \frac{b}{t}\right) w = 0,$$

and so the vectors

$$\frac{t}{\sqrt{1+t^2}} \begin{pmatrix} 0 \\ u - \left(a + \frac{b}{t}\right) v \end{pmatrix},\ \frac{1}{\sqrt{1+t^2}} \begin{pmatrix} q - (a - bt)w \\ 0 \end{pmatrix}$$

are in L_t. These two vectors span a rectangular sublattice R_t of L_t, whose determinant is

$$\det R_t = \left| \frac{t}{1+t^2} \left(u - \left(a + \frac{b}{t} \right) v \right) (q - (a - bt)w) \right| = \frac{b^2 vw(t^2+1)}{|t|}.$$

Let Γ' be a rectangular sublattice of Γ_t isometric to R_t, then

$$[\Gamma_t : \Gamma'] = [L_t : R_t] = \frac{\det R_t}{\det L_t} = \frac{|b| vw(t^2+1)}{|t|}.$$

Now, let E' be the elliptic curve (up to isomorphism) corresponding to Γ' (up to similarity), then the degree of the isogeny $E' \to E_\tau$ is

$$\delta(E'/E_\tau) = [\Gamma_\tau : \Gamma'] = \frac{|b| vw(t^2+1)}{|t|}.$$

This completes the proof. □

Notice that if a lattice Γ_τ for $\tau = a + bi \in \mathcal{D}$ contains a rectangular sublattice, then it contains infinitely many non-similar rectangular sublattices: these can be obtained for instance by multiplying the original rectangular sublattice by matrices of the form $\begin{pmatrix} l & 0 \\ 0 & m \end{pmatrix}$ for relatively prime integers l, m. However all of these sublattices are parallel to each other, meaning that they are spanned by parallel pairs of orthogonal basis vectors. Can Γ_τ have non-parallel rectangular sublattices? This condition is equivalent to saying that there are multiple ways to rotate Γ_τ so that some sublattice will have an orthogonal basis along the coordinate axes. Since the parameter t of Theorem 1.4 is the tangent of the angle of rotation, this can be possible if and only if there exist distinct $t_1, t_2 \in \mathbb{R}$ satisfying condition (1) of Theorem 1.4 so that $t_1 \neq -1/t_2$ (t and $-1/t$ correspond to rotations resulting in the same lattice). This turns out to be possible if and only if τ is a quadratic irrationality, in which case the corresponding elliptic curve E_τ is said to be a curve with *complex multiplication (CM)*: this is precisely the situation when the endomorphism ring of E_τ is larger than $\mathbb{Z}$ (specifically, an order in the imaginary quadratic field $\mathbb{Q}(\tau)$; see Corollary III.9.4 of [14]). We now prove that for such τ there are infinitely many different real numbers t satisfying condition (1) of Theorem 1.4.

Proposition 4.2. *With notation as in Theorem 1.4, suppose that there exist $t_1, t_2 \in \mathbb{R}$ satisfying condition (1). Define*

$$\alpha_1 := a - bt_1 \in \mathbb{Q} \tag{16}$$

$$\beta_1 := a + b/t_1 \in \mathbb{Q} \tag{17}$$

$$\alpha_2 := a - bt_2 \in \mathbb{Q} \tag{18}$$

$$\beta_2 := a + b/t_2 \in \mathbb{Q} \tag{19}$$

Assume also that $t_1 \neq t_2, -1/t_2$. Then $t_1^2, t_2^2 \in \mathbb{Q}$ and $b^2 \in \mathbb{Q}$, meaning that $\tau = a + bi$ is a quadratic irrationality, and hence E_τ is a CM elliptic curve.

Proof. Subtract (16) from (18) and (17) from (19) to obtain

$$b(t_1 - t_2) = \alpha_1 - \alpha_2 \tag{20}$$

$$b\frac{t_1 - t_2}{t_1 t_2} = \beta_1 - \beta_2. \tag{21}$$

Divide (20) by (21) (note that $t_1 \neq t_2$ implies that $\beta_1 \neq \beta_2$) to conclude that

$$t_1 t_2 = \frac{\alpha_1 - \alpha_2}{\beta_1 - \beta_2} \in \mathbb{Q}.$$

Multiply (19) by $t_1 t_2$ and take into the account $bt_1 = a - \alpha_1$ (from (16)) to obtain

$$a(t_1 t_2 + 1) = \alpha_1 + \beta_2 t_1 t_2$$

and conclude that $a \in \mathbb{Q}$ since $t_1 t_2 \neq -1$ by assumption, and the quantities $\alpha_1, \beta_2 \in \mathbb{Q}$, and $t_1 t_2 \in \mathbb{Q}$ are already known to be rational. Since $a \in \mathbb{Q}$, we conclude from (16) and (18) that $bt_1, bt_2 \in \mathbb{Q}$, and therefore their ratio (note that $t_1 t_2 \neq 0$ because $\tau \notin \mathbb{R}$) is rational: $t_1/t_2 \in \mathbb{Q}$. Multiplication (and division) by the rational quantity $t_1 t_2$ now allows us to conclude that $t_1^2, t_2^2 \in \mathbb{Q}$. Finally, since $bt_1 = a - \alpha_1 \in \mathbb{Q}$, we square it to conclude that $b^2 t_1^2 \in \mathbb{Q}$, and divide by t_1^2 to obtain that $b^2 \in \mathbb{Q}$ as claimed. □

Proposition 4.2 asserts essential uniqueness of the real number t satisfying condition (1) of Theorem 1.4 in the generic (non-CM) situation. In the case when CM occurs, there is an infinite family of such t.

Corollary 4.3. *With notation as in Proposition 4.2, assume $\tau = a + bi$ with $a, b^2 \in \mathbb{Q}$. Then every t satisfying condition (1) of Theorem 1.4 is of the form*

$$qb \text{ or } q/b \text{ for some } q \in \mathbb{Q}. \tag{22}$$

Conversely, for every rational q, $t = qb$ satisfies condition (1) of Theorem 1.4.

Proof. Assume t_1, t_2 are two different values of t satisfying condition (1) of Theorem 1.4. From the proof of Proposition 4.2, we know that $t_1/t_2 \in \mathbb{Q}$, so $t_1 = qt_2$ for some $q \in \mathbb{Q}$. Then t_1 satisfies (22) if and only if t_2 does, so it is enough to show that there exists a t of the form (22) satisfying condition (1) of Theorem 1.4. Indeed, for any $q \in \mathbb{Q}$, take $t = qb$, then

$$a - bt = a - b^2 q \in \mathbb{Q}, \ a + b/t = a + 1/q \in \mathbb{Q}.$$

On the other hand, take $t = q/b$, then

$$a - bt = a - q \in \mathbb{Q}, \ a + b/t = a + b^2/q \in \mathbb{Q}.$$

This completes the proof. □

Let us now provide an interpretation of this result for elliptic curves. For every $N > 1$ there is a symmetric polynomial $F_N(X, Y) = F_N(Y, X)$ in two variables with integer coefficients, which has the following property: two elliptic curves E_1 and E_2 with corresponding j-invariants j_1 and j_2 are isogenous with an isogeny of degree N if and only if $F_N(j_1, j_2) = 0$. The polynomial $F_N(X, Y)$ is commonly referred to as *N-th modular polynomial*. Our Theorem 1.4 now implies that an elliptic curve E with a non-real j-invariant $j(E)$ is virtually rectangular if and only if for some N the polynomial $F_N(X, j(E))$, now monic with complex coefficients, has a real root. Another observation that is not difficult to prove is that the curve E_τ is virtually rectangular if and only if $E_{-\bar{\tau}}$ is virtually rectangular, and if this is the case then the curves E_τ and $E_{-\bar{\tau}}$ are isogenous (i.e., any one of the corresponding lattices is similar to a sublattice of the other one).

Further, consider an elliptic curve E over $\mathbb{C}$ with a non-real j-invariant $j(E)$ which is not isomorphic to an elliptic curve over $\mathbb{R}$. Assume that E is nevertheless isogenous to an elliptic curve E' over $\mathbb{R}$ with j-invariant $j(E')$. This implies that

$j(E)$ and $j(E')$ are algebraically dependent over $\mathbb{Q}$. Then the upper bound on the inequality (10) on Theorem 1.4 gives a bound on the degree of the field extension $[\mathbb{Q}(j(E), j(E')) : \mathbb{Q}(j(E))]$.

5. Planar virtually rectangular lattices on the modular curve

Our goal here is to present a geometric interpretation which justifies the consideration of virtually rectangular lattices as very natural objects. Specifically, we look at how the points corresponding to the virtually rectangular lattices are positioned in the moduli space of all lattices.

WR lattices can be clearly seen in the fundamental domain (see Lemma 4.1 above), as can be the rectangular ones: they correspond to the set $\{it : t \in \mathbb{R}, t \geq 1\}$. Virtually WR lattices (those that contain a finite-index WR sublattice) are the same as virtually rectangular in $\mathbb{R}^2$ (see Lemma 2.1 of [11]), however they are not as easily identified in the fundamental domain. Indeed, for a rational $a \in \mathbb{Q}$ with $-1/2 < a \leq 1/2$, the lattice corresponding to every point $\tau = a + ib \in \mathcal{D}$ in the fundamental domain is virtually rectangular. While this set is already too large to be contained on any continuous path in the upper half-plane, there are many more points in the fundamental domain corresponding to virtually rectangular lattices. Surprisingly, the picture becomes much clearer if we look at these points as points on the modular curve instead.

Let

$$\mathfrak{H} = \{\tau = x + iy \in \mathbb{C} \mid y = \Im(\tau) > 0\}$$

be the complex upper half-plane. It comes with the Poincaré metric

$$ds^2 = y^{-2}(dx^2 + dy^2),$$

which is invariant under the action of $\mathrm{PGL}_2(\mathbb{R})^+$ (and is uniquely defined by that property up to a constant multiplication). Let $Y = \mathrm{PSL}_2(\mathbb{Z})\backslash\mathfrak{H}$. The points of Y classify elliptic curves up to isomorphism over $\mathbb{C}$ and correspond to (orientation preserving) similarity classes of lattices $\Gamma_\tau = \langle 1, \tau\rangle_{\mathbb{Z}}$ in the plane.

Since $\mathrm{PSL}_2(\mathbb{Z}) \subset \mathrm{PGL}_2(\mathbb{R})$, the space Y inherits the metric from $\mathfrak{H}$, in particular, geodesics on Y are precisely the images of the geodesics on $\mathfrak{H}$ under the natural projection $\pi : \mathfrak{H} \to Y$ (recall that geodesics for a given metric are the paths of shortest length). The modular curve is the compact Riemann surface $X = \mathrm{PSL}_2(\mathbb{Z})\backslash\overline{\mathfrak{H}}$, where $\overline{\mathfrak{H}} = \mathfrak{H} \cup \mathbb{Q} \cup \{\infty\}$. We then have $X = Y \cup \infty$, and the point at infinity ∞ does not correspond to any lattice.

Geodesics on $\mathfrak{H}$ (for the metric ds^2) are the vertical lines together with the semicircles orthogonal to the real axis (see e.g. [5, Proposition 4.5.5], [12, Lemma 1.4.1]). While all geodesics on $\mathfrak{H}$ are of infinite length, under the map π, some geodesics become closed, while others are still of infinite length. We say that a geodesic on $\mathfrak{H}$ *passes through* ∞ if it is either a vertical line or a semicircle which meets the real line at a rational point, and we say that a geodesic is *closed at* ∞ if it is either vertical with a rational x-coordinate or both ends of the semicircle meet the real line at rational points. We apply the same terminology to the images of these geodesics under the projection π. This terminology is not standard. For example, while for any two points in Y there exists exactly one geodesic which passes through these two points, there are infinitely many geodesics which pass through ∞ and any given point on Y; furthermore, the geodesics which are closed

at ∞ are never closed in Y. However, one may possibly argue that, for example, a semicircle which meets the real line at two rational points is really closed at ∞ because both of its ends map to ∞ under π.

Theorem 5.1. *A point on Y corresponds to a virtually rectangular lattice if and only if this point belongs to a geodesic that is closed at ∞.*

Proof. Let $p \in Y$, and assume that the corresponding lattice is virtually rectangular. Thus $p = \pi(\tau)$ with $\tau = a + bi \in \mathfrak{H}$, and the lattice Γ_τ is virtually rectangular. Then the lattice Γ_τ has two orthogonal vectors, say, $A\tau + B$ and $C\tau + D$ where A, B, C, D are integers and $AD - BC \neq 0$. The orthogonality condition can be written as

$$(a^2 + b^2)AC + a(AD + BC) + BD = 0. \tag{23}$$

If $A = 0$, then $BC \neq 0$, and τ belongs to the vertical line $x = -D/C$. If $C = 0$, then $AD \neq 0$, and τ belongs to the vertical line $x = -B/A$. Finally, if $AC \neq 0$ then (23) becomes an equation of the semicircle

$$\left(a + \frac{AD + BC}{2AC}\right)^2 + b^2 = \frac{(AD - BC)^2}{4A^2C^2}$$

satisfied by (a, b) with a rational radius of $|(AD - BC)/2AC|$ and the center on real line at $x = -(AD + BC)/2AC \in \mathbb{Q}$. Thus in either case $\pi(\tau)$ belongs to a geodesic that is closed at ∞.

Conversely, assume that $\pi(\tau)$ belongs to a geodesic which is closed at ∞. If this geodesic is the image under π of a vertical line in $\mathfrak{H}$ with a rational x-coordinate, then $\tau = a + bi$ with $a \in \mathbb{Q}$, and the lattice Γ_τ is virtually rectangular. Otherwise, τ belongs to a semicircle which meets the real line at rational points, say α and β. Pick $\sigma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\sigma(\alpha) = \infty$. Then $\sigma(\beta) \in \mathbb{Q}$, and σ takes the semicircle to the vertical line with a rational x-coordinate of $\sigma(\beta)$ (Möbius transformations preserve the Poincaré metric, therefore take geodesics to geodesics). Thus the lattice $\langle 1, \sigma(\tau)\rangle_{\mathbb{Z}}$ is virtually rectangular, and this lattice is similar to Γ_τ. We thus conclude that Γ_τ is virtually rectangular. □

Acknowledgement: We wish to thank the anonymous referees for their helpful suggestions, which improved the quality of presentation.

References

[1] I. Aliev, G. Averkov, J. De Loera and T. Oertel, Optimizing sparsity over lattices and semigroups. *Lecture Notes in Computer Science*, 2020.
[2] I. Aliev, J. De Loera, T. Oertel and C. O'Neill, Sparse solutions of linear diophantine equations. *SIAM Journal on Applied Algebra and Geometry*, 1(1), 239–253, 2017.
[3] R. Baker and D. Masser, Siegel's lemma is sharp for almost all linear systems. *Bull. Lond. Math. Soc.*, 51(5), 853–867, 2019.
[4] E. Bombieri and J. D. Vaaler, On Siegel's lemma. *Invent. Math.*, 73(1), 11–32, 1983.
[5] H. Cohen and F. Strömberg, Modular Forms. A Classical Approach. *Graduate Studies in Mathematics, 179.* American Mathematical Society, Providence, RI, 2017.
[6] Y. C. Eldar and G. Kutyniok, Compressed sensing: theory and applications. Cambridge University Press, 2012.
[7] L. Fukshansky, D. Needell and B. Sudakov. An algebraic perspective on integer sparse recovery. *Appl. Math. Comput.*, 340, 31–42, 2019.

[8] L. Fukshansky, P. Guerzhoy and F. Luca. On arithmetic lattices in the plane. *Proc. Amer. Math. Soc.*, 145(4), 1453–1465, 2017.
[9] P. M. Gruber and C. G. Lekkerkerker, Geometry of Numbers. North-Holland Publishing Co., 1987.
[10] M. Koecher and A. Krieg, Elliptische Funktionen und Modulformen. Springer-Verlag, Berlin, 1998.
[11] S. Kühnlein, Well-rounded sublattices. *Int. J. Number Theory*, 8(5), 1133–1144, 2012.
[12] T. Miyake, Modular Forms *Modular forms*, Translated from the 1976 Japanese original by Yoshitaka Maeda. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2006.
[13] W. M. Schmidt, Diophantine approximations and Diophantine equations. Lecture Notes in Mathematics, 1467. Springer-Verlag, Berlin, 1991.
[14] J. H. Silverman, The arithmetic of elliptic curves. Graduate Texts in Mathematics, 106. Springer-Verlag, New York, 1986.

Department of Mathematics, 850 Columbia Avenue, Claremont McKenna College, Claremont, CA 91711
Email address: lenny@cmc.edu

Department of Mathematics, University of Hawaii, 2565 McCarthy Mall, Honolulu, HI, 96822-2273
Email address: pavel@math.hawaii.edu

Institut für Algebra und Geometrie, Fakultät für Mathematik, KIT, FRG-76128 Karlsruhe
Email address: stefan.kuehnlein@kit.edu