Reduction theory and heights of rational points

Joint with Jack Thorne

Jef Laga

University of Cambridge

Motivating question

How complicated are rational points on a curve or an abelian variety?

Dem'janenko–Lang–Silverman conjecture

There exists a constant $c_g > 0$ such that for all g-dimensional ppavs $(A, \lambda)/\mathbb{Q}$ and points $P \in A(\mathbb{Q})$ such that $\mathbb{Z} \cdot P$ is Zariski dense, we have

$$\hat{h}(P) \geq c_g \cdot h(A, \lambda).$$

Consequence of Vojta's conjectures (Ih 2002)

If $\mathcal{C} \to B$ is a family of curves of genus $g \geq 2$, then there are constants c_1, c_2 such that for all $b \in B(\mathbb{Q})$ and $P \in \mathcal{C}_b(\mathbb{Q})$,

$$h(P) \leq c_1 \cdot h(b) + c_2.$$

Given a polynomial $f(x) = x^{2g+1} + c_2 x^{2g-1} + \cdots + c_{2g}x + c_{2g+1} \in \mathbb{Z}[x]$ with $\mathrm{disc}(f) \neq 0$, let

- $\mathrm{Ht}(f) = \max |c_i|^{1/i}$;
- $C_f^0 \colon y^2 = f(x)$ affine curve;
- C_f: projective completion of C_f^0, a genus-g hyperelliptic curve with unique point $P_\infty \in C_f(\mathbb{Q})$ at infinity;
- J_f: the Jacobian variety of C_f.

Expectation

When ordered by height $\mathrm{Ht}(f)$, 50% of J_f have rank 0, 50% of J_f have rank 1, and when $g \geq 2$, 100% of C_f have $C_f(\mathbb{Q}) = \{P_\infty\}$.

When nontrivial points in $J_f(\mathbb{Q})$ do exist, how large are they typically?

If $D = [\sum_{i=1}^{m} P_i - mP_\infty]$ with $P_i \in C_f^0(\bar{\mathbb{Q}})$ and m minimal, let $h^\dagger(D) = \sum h(x(P_i))$, where $h(\cdots)$ denotes the logarithmic Weil height.

Theorem (L.-Thorne, 2024)

Fix $\epsilon > 0$. Then for 100% of $f(x)$ (ordered by height), every nonzero $D \in J_f(\mathbb{Q})$ satisfies

$$h^\dagger(D) \geq (g - \epsilon) \log \mathrm{Ht}(f).$$

Theorem (L.-Thorne, 2025)

Fix $\epsilon > 0$. Then for 100% of $f(x)$ (ordered by height), every nonzero $D \in J_f(\mathbb{Q})$ satisfies

$$\hat{h}(D) \geq \left(\frac{3g-1}{2} - \epsilon\right) \log \mathrm{Ht}(f).$$

Theorem (L.-Thorne)

For 100% of $f(x)$ (ordered by height), every nonzero $D \in J_f(\mathbb{Q})$ satisfies

$$h^{\dagger}(D) \geq (g-\epsilon)\log \mathrm{Ht}(f), \quad \hat{h}(D) \geq \left(\frac{3g-1}{2}-\epsilon\right)\log \mathrm{Ht}(f).$$

Remarks:

- 'density-1 version' of the Dem'janenko–Lang–Silverman conjecture.
- It implies J_f and C_f typically have no 'small height points'.
- In a different paper, we have an analogue of the first version of the theorem for the family of non-monic curves $y^2 = f_0x^{2g+2} + f_1x^{2g+1}z + \cdots + f_{2g+2}z^{2g+2} \in \mathbb{Z}[x,z]$ and points of odd degree.

Theorem (L.-Thorne)

For 100% of $f(x)$ (ordered by height), every nonzero $D \in J_f(\mathbb{Q})$ satisfies

$$h^{\dagger}(D) \geq (g-\epsilon)\log \mathrm{Ht}(f), \quad \hat{h}(D) \geq \left(\frac{3g-1}{2} - \epsilon\right)\log \mathrm{Ht}(f).$$

Proof strategy:

1. Define a different 'height' $\tilde{h}\colon J_f(\mathbb{Q}) \to \mathbb{R}$, in terms of reduction theory.
2. Show that $\tilde{h}(D) \geq -\epsilon \log \mathrm{Ht}(f)$ for all nonzero D in a density 1 family.
3. Relate $\tilde{h}$ to $h^{\dagger}$ and $\hat{h}$.

My focus: Explaining Steps 1 and 2.

Jack's talk: Relating $\tilde{h}$ and $\hat{h}$, and much more!

The goal of reduction theory

Given an action of a group Γ on a set S, find representatives that are 'small' or 'reduced'.

Example 1

$\Gamma = \mathrm{SL}_2(\mathbb{Z})$, acting on
$S = \{\text{Positive definite } Q(x, y) = ax^2 + bxy + cy^2 \in \mathbb{R}[x, y]\}$.

Example:

$$458x^2 + 214xy + 25y^2 \rightsquigarrow x^2 + y^2$$

Reduction algorithm:

1. Let $\tau =$ unique root of $Q(x, 1)$ with $\mathrm{Im}(\tau) > 0$.
2. Let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma \cdot \tau \in \mathcal{F}$, the Minkowski fundamental domain.
3. Then $\gamma \cdot Q$ is 'reduced', in the sense that $|b| \leq a \leq c$

Example 2: lattice reduction

Given a lattice $(\Lambda, \langle -, - \rangle)$, find a 'small' $\mathbb{Z}$-basis of Λ.
Equivalently: given an inner product H on $\mathbb{R}^n$, find a $\gamma \in \mathrm{SL}_n(\mathbb{Z})$ such that $\gamma H \gamma^t$ has small coefficients.

LLL algorithm: efficient algorithm to find 'LLL-reduced' representative.
Example: running LLL on

$$H = \begin{pmatrix} 176413988.185 & -11560848.1174 & 3471.84429193 \\ -11560848.1174 & 757736.524016 & -1499.92503970 \\ 3471.84429193 & -1499.92503970 & 13237.5156939 \end{pmatrix}$$

gives

$$\gamma = \begin{pmatrix} 0 & 0 & 1 \\ 4 & 61 & 6 \\ -3 & -46 & -4 \end{pmatrix}, \quad \gamma H \gamma^t = \begin{pmatrix} 13237.5 & 1817.04 & 5630.96 \\ 1817.04 & 12789.5 & -1067.59 \\ 5630.96 & -1067.59 & 45450.2 \end{pmatrix}.$$

Example 3 (Cremona–Stoll)

$\Gamma = \mathrm{SL}_2(\mathbb{Z})$, acting on
$S = \{$Binary n-ic forms $f(x,y) \in \mathbb{R}[x,y]$ with $\mathrm{disc}(f) \neq 0\}$.
Every f has a 'Julia covariant' $Q_f \in \mathbb{R}[x,y]_{\deg=2}$, and $Q_{\gamma\cdot f} = \gamma \cdot Q_f$.
Reducing $f \Leftrightarrow$ reducing Q_f.

Example 4 (Cremona–Fisher–Stoll)

$\mathrm{SL}_3(\mathbb{Z})$ acting on
$\{$Ternary cubic forms $f(x,y,z) \in \mathbb{R}[x,y,z]$ with $\mathrm{disc}(f) \neq 0\}$.
Every f has an $\mathrm{SL}_3(\mathbb{R})$-covariant inner product H_f on $\mathbb{R}^3$.
Reducing $f \Leftrightarrow$ reducing H_f.

For example, running this on

$$f = 40877301x^3 - 11504y^3 + 12z^3 - 8035425x^2y - 64887x^2z \\ +526580xy^2 - 200y^2z + 5803xz^2 - 383yz^2 + 7307xyz$$

gives a $\gamma \in \mathrm{SL}_3(\mathbb{Z})$ such that

$$\gamma \cdot f = 12x^3 + 12y^3 + 171z^3 + 65x^2y + 65x^2z$$

In all these examples, we have a GL_n-representation V, an open subset $V^s \subset V$ and a $\mathrm{GL}_n(\mathbb{R})$-equivariant map

$$\mathcal{R}\colon V^s(\mathbb{R}) \to \{\text{Inner products on } \mathbb{R}^n\}.$$

Reducing an element $v \in V^s(\mathbb{R})$ boils down to reducing $\mathcal{R}(v)$.

Question

Can we generalize this to groups other than GL_n?

First step

$\gamma \mapsto \gamma\gamma^t$ identifies $\mathrm{GL}_n(\mathbb{R})/\mathrm{O}_n(\mathbb{R})$ with $\{\text{Inner products on } \mathbb{R}^n\}$.

Given a reductive group G acting on V, is there a $G(\mathbb{R})$-equivariant map

$$\mathcal{R}\colon V^s(\mathbb{R}) \to X_G, \quad (\text{'reduction covariant'})$$

where $X_G = G(\mathbb{R})/K_\infty$, such that reduction theory works similarly?

Sometimes, yes

Such $\mathcal{R}$ exists for every stable Vinberg representation (G, V) (Thorne), and there is an analogue of LLL for X_G for arbitary $G/\mathbb{Z}$ (Thorne–Romano).

Let $g \geq 1$ be an integer, $W = \mathbb{Z}^{2g+1}$ and J the bilinear form with Gram matrix

$$J = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}.$$

Let

$$G = \mathrm{SO}_J \leq \mathsf{GL}(W)$$

and

$$V = \{T \in \mathsf{End}(W) \colon T^* = T, \mathrm{Tr}(T) = 0\}.$$

G acts on V via $g \cdot T = gTg^{-1}$.
Each $T \in V$ has a characteristic polynomial $f_T = \det(xI - T)$.
Let $V^s \subset V$ be the subset such that f_T has distinct roots.
We will define a reduction covariant

$$\mathcal{R} \colon V^s(\mathbb{R}) \to X_G$$

$$G = \mathrm{SO}_J \curvearrowright V = \{T \in \mathsf{End}(W) \colon T^* = T, \mathrm{Tr}(T) = 0\}$$

The map $\gamma \mapsto \gamma\gamma^t$ identifies X_G with the set of inner products H on $\mathbb{R}^{2g+1}$ compatible with J, in the sense that $J = HJH$.

Lemma

If $T \in V^s(\mathbb{R})$, there exists a unique inner product H_T on W satisfying:

1. H_T is compatible with J; and
2. T commutes with its H_T-adjoint.

We may define

$$\mathcal{R} \colon V^s(\mathbb{R}) \to X_G$$

by $\mathcal{R}(T) = H_T$. This is our 'reduction covariant'.

Example in $g = 3$ (Thorne):

$$T = \begin{pmatrix} -14 & 1 & 0 & 0 & 0 & 0 & 0 \\ -195 & 0 & 1 & 0 & 0 & 0 & 0 \\ -2728 & 0 & 7 & 0 & -1 & 0 & 0 \\ -10237 & 0 & 0 & 14 & 0 & 0 & 0 \\ 19095 & -6 & -48 & 0 & 7 & 1 & 0 \\ 1546 & -26 & -6 & 0 & 0 & 0 & 1 \\ 390 & 1546 & 19095 & -10237 & -2728 & -195 & -14 \end{pmatrix} \in V(\mathbb{Z})$$

gives

$$\mathcal{R}(T) = \begin{pmatrix} 3.74708 & 53.7691 & 750.242 & 2813.43 & -5244.78 & -421.526 & -47.2448 \\ 53.7691 & 776.143 & 10830.1 & 40612.6 & -75708.6 & -6080.03 & -681.676 \\ 750.242 & 10830.1 & 151130. & 566729. & -1.05648 \times 10^6 & -84842.6 & -9520.71 \\ 2813.43 & 40612.6 & 566729. & 2.12521 \times 10^6 & -3.96175 \times 10^6 & -318157. & -35704.6 \\ -5244.78 & -75708.6 & -1.05648 \times 10^6 & -3.96175 \times 10^6 & 7.38537 \times 10^6 & 593097. & 66564.2 \\ -421.526 & -6080.03 & -84842.6 & -318157. & 593097. & 47660.8 & 5338.34 \\ -47.2448 & -681.676 & -9520.71 & -35704.6 & 66564.2 & 5338.34 & 660.273 \end{pmatrix}.$$

Applying an LLL-type algorithm to $\mathcal{R}(T) \in X_G$ shows T is $G(\mathbb{Z})$-equivalent to

$$T' = \begin{pmatrix} 0 & 0 & -1 & 2 & 2 & -2 & 3 \\ 1 & 0 & 1 & 0 & 0 & 0 & -2 \\ 0 & 1 & 1 & -2 & 0 & 0 & 2 \\ 0 & 0 & 1 & -2 & -2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

For every $f = x^{2g+1} + c_2 x^{2g-1} + \cdots \in \mathbb{Z}[x]$ with $\mathrm{disc}(f) \neq 0$, Bhargava–Gross (2012) constructed a map

$$J_f(\mathbb{Q})/2J_f(\mathbb{Q}) \hookrightarrow G(\mathbb{Q}) \setminus V_f(\mathbb{Q}),$$

where $V_f = \{T \in V : f_T = f\}$.

Building on their work, we lift this map to a map

$$\eta_f : J_f(\mathbb{Q}) \to G(\mathbb{Z}) \setminus V_f(\mathbb{Z}).$$

Rough idea

Let $D = [E - mP_\infty] \in J_f(\mathbb{Q})$ and $\mathcal{W} : (y = 0) \subset C_f^0$.
Let $W_D = \mathrm{H}^0(\mathcal{W}, \mathcal{O}_{C_f}(E)|_{\mathcal{W}})$. We construct:

- An split symmetric form $(-,-)_D$ on W_D;
- A self-adjoint linear operator $T_D : W_D \to W_D$ with char poly f;
- An integral structure on W_D.

A choice of isomorphism $(W_D, (-,-)_D) \simeq (W, (-,-)_J)$ maps T_D to an element of $V_f(\mathbb{Z})$, well defined up to $G(\mathbb{Z})$-conjugation.

$$J_f(\mathbb{Q}) \xrightarrow{\eta_f} G(\mathbb{Z})\backslash V_f(\mathbb{Z}) \xrightarrow{\mathcal{R}} G(\mathbb{Z})\backslash X_G.$$

Conclusion

Every $D \in J_f(\mathbb{Q})$ determines a rank $2g+1$ lattice Λ_D (with extra data).

Tantalizing question: what is the relation between D and Λ_D?

Proposition

Let $D = [\sum_{i=1}^{m} P_i - mP_\infty]$ with $P_i = (x_i, y_i)$.
Let $U(x) = \prod(x - x_i) \in \mathbb{Q}[x]$ and N be the denominator of $U(x)$.
Assume $y_i \neq 0$ for all i.
Then there exists a primitive vector $v_D \in \Lambda_D$ such that

$$(v_D, v_D) = N \sum_{i=1}^{2g+1} \frac{|U(\omega_i)|}{|f'(\omega_i)|},$$

where $\omega_1, \ldots, \omega_{2g+1} \in \mathbb{C}$ are the roots of $f(x)$.

Proposition

There exists a primitive vector $v_D \in \Lambda_D$ such that

$$(v_D, v_D) = N \sum_{i=1}^{2g+1} \frac{|U(\omega_i)|}{|f'(\omega_i)|},$$

where $U(x) = \prod(x - x(P_i)) \in \mathbb{Q}[x]$ if $D = [\sum_{i=1}^{m} P_i - mP_\infty]$, $\omega_1, \ldots, \omega_{2g+1} \in \mathbb{C}$ are the roots of $f(x)$, and N is the denominator of $U(x)$.

Definition

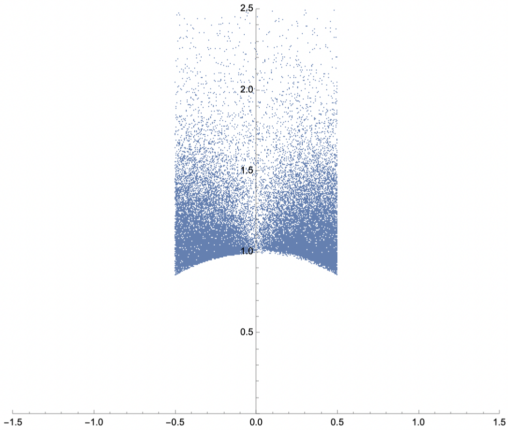
$$\tilde{h}(D) = \frac{1}{2} \log(v_D, v_D).$$

It remains to prove that for 100% of f, every nonzero D satisfies

$$\tilde{h}(D) \geq -\epsilon \log \mathrm{Ht}(f).$$

If $\tilde{h}(D)$ is very negative, then v_D is a very short vector of Λ_D, and Λ_D is very 'skew'. Can this happen often?

Suppose $g = 1$. Then $G \simeq \mathrm{PGL}_2$, and $X_G \simeq$ upper half plane.
Plotting the elements $\mathrm{PGL}_2(\mathbb{Z}) \cdot \mathcal{R}(v) \in G(\mathbb{Z}) \setminus X_G$ for many small v looks like this:



This suggests equidistribution!

Using geometry-of-numbers techniques, we show

Theorem

For every $g \geq 1$ *and as* $\mathrm{Ht}(f_T) \to +\infty$, *the map*

$$\mathcal{R}\colon G(\mathbb{Z}) \setminus V(\mathbb{Z})^{\mathrm{irr}} \to G(\mathbb{Z})\backslash X_G$$

equidistributes with respect to the natural probability measure on $G(\mathbb{Z})\backslash X_G$.

(Here $v \in V(\mathbb{Z})$ is irreducible if it is not $G(\mathbb{Q})$-conjugate to $\eta_f(0)$.)

Punchline

In the moduli space $G(\mathbb{Z})\backslash X_G$, most lattices do not have very short vectors!

More precisely, the subset $U_\delta \subset G(\mathbb{Z})\backslash X_G$ of lattices Λ such that there is a $v \in \Lambda$ with $0 \neq (v, v) \leq \delta$ has $\mu(U_\delta) \to 0$ as $\delta \to 0$.

Theorem

$$\mathcal{R}\colon G(\mathbb{Z}) \setminus V(\mathbb{Z})^{\mathrm{irr}} \to G(\mathbb{Z})\backslash X_G$$

equidistributes with respect to the natural probability measure on $G(\mathbb{Z})\backslash X_G$.

Proof that $\tilde{h}(D) \geq -\epsilon \log \mathrm{Ht}(f)$ for 100% of f, assuming $D \notin 2J_f(\mathbb{Q})$:

- Suppose $\tilde{h}(D) < -\epsilon \log \mathrm{Ht}(f)$ for some $D \in J_f(\mathbb{Q})$ for a positive proportion c of f;
- Then, for every $\delta > 0$, a fixed positive proportion c of f have the property that there exists an element $T \in G(\mathbb{Z})\backslash V_f(\mathbb{Z})^{\mathrm{irr}}$ such that $\Lambda = G(\mathbb{Z}) \cdot \mathcal{R}(T)$ has a vector v with $0 \neq (v, v) \leq \delta$.
- Therefore a fixed positive proportion c' of irreducible $G(\mathbb{Z})$-orbits in $V(\mathbb{Z})$ have reduction covariant with a vector v of norm $\leq \delta$.
- Taking δ so that $\mu(U_\delta) < c'$, together with the equidistribution theorem, gives a contradiction.

An additional argument handles 2-divisible points.

Summary:

1. For a certain representation V of $G = \mathrm{SO}_{2g+1}$, we define a reduction covariant $\mathcal{R}\colon V^s(\mathbb{R}) \to X_G$.
2. For every f, we define a map $\eta\colon J_f(\mathbb{Q}) \to G(\mathbb{Z}) \setminus V(\mathbb{Z})$.
3. For every $D \in J_f(\mathbb{Q})$, we get a rank $2g+1$ lattice $\Lambda_D = \mathcal{R}(\eta(D)) \in G(\mathbb{Z}) \setminus X_G$.
4. We find a vector $v_D \in \Lambda_D$ such that $\log ||v_D||$ can be related to height functions like $h^\dagger(D)$
5. If a positive proportion of f have a $D \in J_f(\mathbb{Q})$ of small height, then many Λ_D have a small vector.
6. This contradicts the equidistribution of $\mathcal{R}$ and the fact that most lattices in X_G do not have a very small vector.