Quadraticity of height functions

Padmavathi Srinivasan

Week 6

Last time we defined the Weil height function of an elliptic curve E defined over a number field K as the function

$$\begin{aligned} h_E\colon E(\overline{\mathbb{Q}}) &\to \mathbb{R} \\ P &\mapsto h(x(P)) \end{aligned}$$

We reduced the proof of the descent step of deducing the Mordell-Weil theorem from its weak version to the following almost parallelogram law for the function h_E.

Theorem 1. [Sil09, Chapter 8, Theorem 6.2] *Let E be an elliptic curve over a number field K. Then for all $P, Q \in E(\overline{\mathbb{Q}})$, we have*

$$h_E(P+Q) + h_E(P-Q) = 2h_E(P) + 2h_E(Q) + O(1), \tag{1}$$

where the implied constants in $O(1)$ depend on E, but are independent of the points P, Q.

For the proof of the Weak Mordell-Weil theorem, we refer the reader to [Sil09, Chapter 8, Section 1]. Today we will prove the almost parallelogram law. The proof involves some explicit algebra using formulas for the group law of the elliptic curve. The third main feature of height functions (after the Northcott property and the existence of local decompositions) that is crucial in this proof is the

"functoriality of heights under morphisms of projective spaces."

1 Functoriality of heights and the Weil height machine

Proposition 2. [Sil09, Chapter 8, Theorem 5.6] *Suppose $F\colon \mathbb{P}^N \to \mathbb{P}^M$ is a morphism of degree d over a number field K, i.e.*

$$F(P) = [f_0(P)\colon \ldots \colon f_M(P)],$$

where the f_i are homogeneous polynomials of degree d in $N+1$ variables with coefficients in the field K. Assume that the f_i have no common zeroes in $\overline{\mathbb{Q}}^{N+1} \setminus (0, 0, \ldots, 0)$. Then there are constants C_1, C_2 depending only on F and not on P such that

$$dh(P) + C_1 \leq h(F(P)) \leq dh(P) + C_2.$$

Proof. Let M_K be the collection of normalized absolute values on a number field K[1], and define

$$n_v := \begin{cases} 2 & \text{if } v \text{ is an Archimedean place corresponding to a pair of complex conjugate embeddings} \\ 1 & \text{otherwise.} \end{cases}$$

Recall that if $P = [x_0 \colon \ldots : x_N] \in \mathbb{P}^N(K)$, then

$$h(P) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log \max(|x_0|_v^{n_v}, \ldots, |x_N|_v^{n_v}).$$

Using the triangle inequality and multiplicativity of absolute values, we will show that for each $v \in M_K$, we have

$$\log \max(|f_0(P)|_v, \ldots, |f_M(P)|_v) = d \log \max(|x_0|_v, \ldots, |x_N|_v) + O(1),$$

where the implied constants in $O(1)$ depend only the morphism F, and not on the place v or the point P. Define

$$\epsilon_v := \begin{cases} 1 & \text{if } v \text{ is Archimedean} \\ 0 & \text{otherwise.} \end{cases}$$

Step 1: Observe that each f_i is a homogeneous polynomial of degree d in $N+1$ variables, and is therefore a sum of at most $\binom{N+d}{d}$ monomials. Fix $v \in M_K$. Let

$$|F|_v = \max\{|a|_v \ : \ a \text{ is a coefficient of some } f_i\}.$$

Multiplicativity and the triangle inequality for absolute values implies that for each i in $\{0, 1, \ldots, M\}$, we have

$$|f_i(P)|_v \leq \binom{N+d}{d}^{\epsilon_v} |F|_v (\max(|x_0|_v, \ldots, |x_N|_v))^d. \tag{2}$$

Step 2: Take the logarithm of both sides of (2), scale by n_v and sum over all places and use $\sum_{v \in M_K} \epsilon_v n_v = [K:\mathbb{Q}]$ to get

$$h(F(P)) \leq dh(P) + C_2,$$

for a constant C_2 that only depends on N, d, F and is independent of P. Observe that Steps 1 and 2 do not need the assumption that the f_i have no common zeroes in $\overline{\mathbb{Q}}^{N+1} \setminus (0, 0, \ldots, 0)$.

Step 3: One way to obtain an inequality in the other direction would be if we could express the coordinates of the input point P, or at least some power of the coordinates of P in terms of the coordinates of $F(P)$. For example, in the case of the morphism

$$\begin{aligned} F \colon \mathbb{P}^1 &\to \mathbb{P}^2 \\ [x_0 : x_1] &\mapsto [x_1^2 - x_0^2 : 2x_0x_1 : x_0^2 + x_1^2] \end{aligned}$$

[1]Our definition of n_v below is different from the one in Silverman's book because we defined normalized our absolute values slightly differently from the book.

parametrizing Pythagorean triples, we can explicitly see that

$$x_0^2 = \frac{1}{2}(f_0(P) + f_2(P))$$
$$x_1^2 = \frac{1}{2}(f_0(P) + f_2(P))$$

The Hilbert Nullstellensatz asserts that this is always possible from the assumption that the f_i have non common zeroes in $\overline{\mathbb{Q}}^{N+1} \setminus (0, 0, \ldots, 0)$, namely that the assumption guarantees that [2] there is an exponent e and that there are polynomials $g_{ij} \in K[x_0, \ldots, x_N]$ for $i \in \{0, 1, \ldots, N\}, j \in \{0, 1, \ldots, M\}$ such that for every i, we have

$$x_i^e = \sum_{j=0}^{M} g_{ij} f_j. \tag{3}$$

Note that the g_{ij} have degree $e - d$. Fix $v \in M_K$. Let

$$|G|_v = \max\{|b|_v \ : \ b \text{ is a coefficient of some } g_{ij}\}.$$

By arguing similarly to Step 1 using multiplicativity and the triangle inequality we get that there is a constant C such that for every i, we have

$$|x_i|_v^e \leq C^{\epsilon_v} \left(\max_{j=0}^{M} |g_{ij}(P)|_v \right) |F(P)|_v$$

Combining this with the inequality

$$|g_{ij}(P)|_v \leq \binom{N+e-d}{e-d}^{\epsilon_v} |G|_v (\max(|x_0|_v, \ldots, |x_N|_v))^{e-d},$$

and taking the maximum over all i yields the inequality

$$|P|_v^e \leq (C')^{\epsilon_v} |P|_v^{e-d} |F(P)|_v,$$

where C' is a constant depending on N, d, F but is independent of the point P. Upon rearranging, we get

$$|P|_v^d \leq (C')^{\epsilon_v} |F(P)|_v. \tag{4}$$

Step 4: Take logs of both sides of (4), scale by n_v and sum over all places and use $\sum_{v \in M_K} \epsilon_v n_v = [K : \mathbb{Q}]$ to get

$$dh(P) + C_1 \leq h(F(P)),$$

for some constant C_1 that depends on N, d, F, but is independent of the point P. □

Suggested exercises 3. Fill in the missing details in the proof above, for example in Steps 2, 3 and 4.

[2] It is clear that if there is an exponent e and polynomials g_{ij} such that (3) holds for every i, then the only common zero of the f_i is $(0, \ldots, 0)$. The Nullstellensatz asserts the converse. The name "Null-stellen-satz" translates to "zero-locus-theorem".

Remark 4. The left inequality in Proposition 2 can fail without the assumption that the f_i have no common zero apart from the origin. For example, the degree 2 rational map

$$\begin{aligned} F\colon \mathbb{P}^2 &\to \mathbb{P}^2 \\ [x:y:z] &\mapsto [x^2 : xy : z^2] \end{aligned}$$

is not defined (i.e. the f_i have a common zero) at exactly one point in $\mathbb{P}^2$, namely $[0:1:0]$. One can show there are infinitely many points P in $\mathbb{P}^2(\mathbb{Q})$ (necessarily of larger and larger height by the Northcott property) such that

$$h(F(P)) = h(P).$$

In particular, there is no constant C_1 such that for all P in $\mathbb{P}^2(\mathbb{Q})$ we have

$$2h(P) + C_1 \leq h(F(P)).$$

Suggested exercises 5. Prove the claim above, namely that there are infinitely many points P in $\mathbb{P}^2(\mathbb{Q})$ such that $h(F(P)) = h(P)$ for the F in Remark 4.

Remark 6. There is a way to intrinsically characterize all embeddings of an abstract variety X into projective space using the notion of line bundles on X. There is even a group structure on the collection of line bundles and the corresponding group is called the Picard group $\mathrm{Pic}(X)$ of X. (The exercise from Week 1 on Segre embeddings is related to how you add two embeddings, and shows that the height function corresponding to the sum of two embeddings is the sum of the two height functions, up to bounded functions.) Weil came up with a way to systematically package all the different height functions on X together into what is now called the Weil height machine. He showed that the various height functions are uniquely determined (up to bounded functions) if we normalize and fix the height on projective space to be the one that we have defined, and further demand that the assignment be functorial in morphisms. More precisely, he showed the following.

Theorem 7. [Sil94, Chapter 3, Section 10, Theorem 10.1, Remark 10.1.1] *For every variety V defined over $\overline{\mathbb{Q}}$, there is a unique homomorphism*

$$h_V\colon \mathrm{Pic}(V) \to \frac{\{\text{functions } V(\overline{\mathbb{Q}}) \to \mathbb{R}\}}{\{\text{bounded functions } V(\overline{\mathbb{Q}}) \to \mathbb{R}\}},$$

such that the height with respect to the line bundle $\mathcal{O}(1)$ on $\mathbb{P}^r$ is the height function we have defined, and such that $h_{V,\psi^* D} = h_{W,D} \circ \psi$ for every morphism $\psi\colon V \to W$.

We have all the ingredients for proving this theorem in these notes and exercises (for example, the one on the Segre embeddings and the functoriality statement we proved above), and essentially done most of the necessary work already. We refer the reader to [Sil94] for details on how to assemble the ingredients together. For an elegant alternate proof of functoriality of heights that does not use the Nullstellensatz, see [Ser89, Section 2.3].

2 The almost parallelogram law for height functions

For the almost parallelogram law, we will also need the following proposition, which is a generalization of the comparison inequality between two different height functions for an algebraic number that we proved in Lecture 2. Let $\alpha_1, \ldots, \alpha_n$ be any n algebraic numbers (not necessarily conjugate). Define

$$f(x) = a_0(x - \alpha_1)\ldots(x - \alpha_n) = a_0x^n + a_1x^{n-1} + \ldots + a_n.$$

Proposition 8. [Sil09, Chapter 8, Theorem 5.9]

$$-n\log(2) + \sum_{i=1}^{n} h(\alpha_i) \leq h([a_0 \colon \ldots \colon a_n]) \leq (n-1)\log(2) + \sum_{i=1}^{n} h(\alpha_i)$$

We leave the proof as an exercise to the reader, but only mention that it involves proving "place by place" inequalities, similar to the proof of Proposition 2.

Corollary 9. *Let $s\colon \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^2$ be the map $s([\alpha_1 : \beta_1], [\alpha_2 : \beta_2]) = [\beta_1\beta_2 : \alpha_1\beta_2 + \alpha_2\beta_1 : \alpha_1\alpha_2]$. Then*

$$h(s(P, Q)) = h(P) + h(Q) + O(1).$$

Proof. If $\beta_1 = 0$ or $\beta_2 = 0$, then $h(s([\alpha_1 : \beta_1], [\alpha_2 : \beta_2])) = h([\alpha_1 : \beta_1]) + h([\alpha_2 : \beta_2])$. So we may assume that $\beta_1 = \beta_2 = 1$ without any loss of generality. Now the corollary follows from Proposition 8 with $n = 2$ since $(x-\alpha_1)(x-\alpha_2) = x^2 - (\alpha_1+\alpha_2)x + \alpha_1\alpha_2 = x^2 + a_1x + a_2$. □

We are now finally ready to prove Theorem 1.

Proof of Theorem 1. Let $G\colon E \times E \to E \times E$ be the morphism defined by $G(P, Q) := (P + Q, P - Q)$, and let $X\colon E \times E \to \mathbb{P}^1 \times \mathbb{P}^1$ be the morphism defined by $X(P, Q) := (x(P), x(Q))$.

The key to this proof is showing that there is a degree 2 morphism $g\colon \mathbb{P}^2 \to \mathbb{P}^2$ that makes the following diagram commute and then applying functoriality of heights to g.

$$\begin{array}{ccc} E \times E & \xrightarrow{\;G\;} & E \times E \\ \downarrow{\scriptstyle X} & & \downarrow{\scriptstyle X} \\ \mathbb{P}^1 \times \mathbb{P}^1 & & \mathbb{P}^1 \times \mathbb{P}^1 \\ \downarrow{\scriptstyle s} & & \downarrow{\scriptstyle s} \\ \mathbb{P}^2 & \xrightarrow{\;g\;} & \mathbb{P}^2 \end{array}$$

We will first show how to conclude the proof assuming the existence of such a g. Let $\sigma = s \circ X$. Using the commutativity of the diagram, we have

$$\begin{aligned} h(\sigma(P + Q, P - Q)) &= h(\sigma \circ G(P, Q)) \\ &= h(g \circ \sigma(P, Q)) \\ &= 2h(\sigma(P, Q)) + O(1), \end{aligned} \tag{5}$$

where the last line follows from Proposition 2 applied to the degree 2 morphism g. Now, Corollary 9 gives

$$h(\sigma(P,Q)) = h(s(x(P),x(Q))) = h(x(P)) + h(x(Q)) + O(1) = h_E(P) + h_E(Q) + O(1), \quad (6)$$

and similarly also gives

$$h(\sigma(P+Q,P-Q)) = h_E(P+Q) + h_E(P-Q) + O(1). \quad (7)$$

Combining (5),(6) and (7) proves the almost parallelogram law, namely that

$$h_E(P+Q) + h_E(P-Q) = 2h_E(P) + 2h_E(Q) + O(1),$$

where the implied constants in the $O(1)$ only depend on the elliptic curve E, and not on the points P, Q. It remains to show the existence of the degree 2 morphism g as above.

Let $y^2 = x^3 + Ax + B$ be the defining equation for the elliptic curve E, where A, B are constants in K such that $4A^3 + 27B^2 \neq 0$. One can show using the explicit description of the group law on E that if $x(P) = [x_1 : 1], x(Q) = [x_2 : 1], x(P+Q) = [x_3 : 1]$ and $x(P-Q) = [x_4 : 1]$ (where $x_i = \infty$ if the corresponding point at infinity on $\mathbb{P}^1$), then

$$\begin{aligned} x_3 + x_4 &= \frac{2(x_1+x_2)(A+x_1x_2)+4B}{(x_1+x_2)^2 - 4x_1x_2} \\ x_3x_4 &= \frac{(x_1x_2 - A)^2 - 4B(x_1+x_2)}{(x_1+x_2)^2 - 4x_1x_2}. \end{aligned} \quad (8)$$

The commutativity of the diagram forces

$$g([1 : x_1 + x_2 : x_1x_2]) = [1 : x_3 + x_4 : x_3x_4],$$

and setting $x_1 + x_2 = (u/t), x_1x_2 = (v/t)$, this in turn forces the formula

$$g([t : u : v]) = [u^2 - 4tv : 2u(At+v) + 4Bt^2 : (v - At)^2 - 4Btu].$$

To show that g is a morphism, we need to show that the only common zero of the three polynomials

$$\begin{aligned} g_0(t,u,v) &:= u^2 - 4tv, \\ g_1(t,u,v) &:= 2u(At+v) + 4Bt^2, \\ g_2(t,u,v) &:= (v - At)^2 - 4Btu \end{aligned} \quad (9)$$

is $(0,0,0)$. This follows from the assumption that $4A^3 + 27B^2 \neq 0$ and exercise 11 below. [3] □

[3] A way to think of the morphism g without writing down explicit formulas is the following. One may think of $\mathbb{P}^2$ as parametrizing unordered pairs of points on $\mathbb{P}^1$, and the map $s\colon \mathbb{P}^1 \times \mathbb{P}^1 \to \mathbb{P}^2$ is the map that "forgets the ordering" by sending a pair of points on $\mathbb{P}^1$ to the coefficients of the polynomial which has the two points as roots. Each unordered pair $x(P), x(Q)$ of points on $\mathbb{P}^1$ corresponds (generically) to 4 points on the elliptic curve E, namely $(P,-P),(Q,-Q)$, which naturally come in pairs. One may think of the map g as the map that sends the collection of 4 points $(P,-P),(Q,-Q)$ on E to the collection of 4 points $(P+Q,-P-Q),(P-Q,Q-P)$ on E obtained by taking all possible pairwise sums of the points in $(P,-P)$ with $(Q,-Q)$.

Suggested exercises 10. Prove the formulas in (8) above.

Suggested exercises 11. Let A, B be elements of K such that $4A^3 + 27B^2 \neq 0$. Let g_0, g_1, g_2 in $K[t, u, v]$ be defined as in (9). Let (t, u, v) be a common zero of g_0, g_1 and g_2.

(a) Show that if $t = 0$, then $u = v = 0$.

(b) Assume $t \neq 0$. Define $z := u/2t$. Using $g_0 = 0$, show that $z^2 = v/t$.

(c) Define $\psi(z) := 4z(A+z^2)+4B$ and $\varphi(z) := (z^2-A)^2-8Bz$. Show that $g_1(t, u, v) = t^2\psi(z)$ and $g_2(t, u, v) = t^2\varphi(z)$.

(d) Verify that $(12z^2 + 16A)\varphi(z) - (3z^3 - 5Az - 27B)\psi(z) = 4(4A^3 + 27B^2)$.

(e) Conclude that ψ and φ cannot simultaneously vanish, and hence g_0, g_1, g_2 have no common zero with $t \neq 0$.

Conclude that if (t, u, v) is a common zero of g_0, g_1 and g_2, then $t = u = v = 0$.

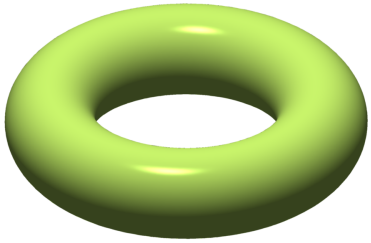
3 The arithmetic of higher genus curves

Having proven the Mordell-Weil theorem, one may ask what is known about the structure of the set of rational points on other classes of varieties. In 1922, Mordell also conjectured that a curve defined over a number field that is sufficiently geometrically complex has only finitely many rational points. More precisely, we have the following landmark theorem of Faltings proving Mordell's conjecture.

Theorem 12. [Fal83] *If X is a smooth projective geometrically connected curve of genus $g \geq 2$ defined over a number field K, then the set $X(K)$ is finite.*

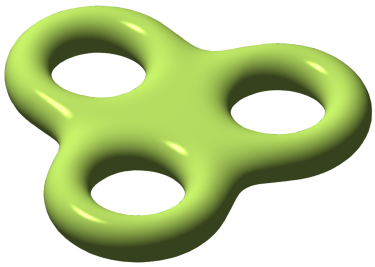
The genus of X in the statement above is a topological invariant of the set of complex points $X(\mathbb{C})$ and is a measure of geometric complexity. The hypothesis that X is smooth projective geometrically connected curve translates to $X(\mathbb{C})$ being a compact connected 1 dimensional complex manifold, that is, a Riemann surface. By the classification theorem of such manifolds, it follows that $X(\mathbb{C})$ is a g-holed torus for some g. An elliptic curve has genus 1, and more generally a hyperelliptic curve with defining equation $y^2 = f(x)$ with f squarefree polynomial of degree $2g + 1$ or $2g + 2$ over $\mathbb{C}$ has genus g.

Although the set of rational points on curves of genus $g \geq 2$ do not have a group structure, there is a canonically associated g-dimensional group variety defined over K called the Jacobian J. If X has a rational point defined over K, then there is an associated Abel-Jacobi map $X \to J$ defined over K. For an elliptic curve, the Abel-Jacobi map is an isomorphism. When K is a number field, the group of rational points $J(K)$ on this g-dimensional group variety is a finitely generated abelian group – this is also called the Mordell-Weil theorem. Key to Faltings's finiteness theorem is a comparison of two different height functions on the space of all Jacobians. See this beautiful survey article [Maz86] by Mazur summarizing the key ideas in the proof of Faltings's theorem without assuming substantial background. Faltings's proof tantalizingly does not give an effective method to



A genus 1 curve

A genus 2 curve



A genus 3 curve

Images from the wikipedia page on genus g Riemann surfaces `https://en.wikipedia.org/wiki/Genus_g_surface`

compute the finite set of points $X(K)$. The quest for alternate effective proofs for Mordell's conjecture continues, and is an active area of research today.

One could alternately ask for bounds on the size of the set $X(K)$. In the same article referenced above, Mazur asks if there is a *uniform* bound on the cardinality of $X(K)$, that depends only on natural invariants associated to X, and not on X itself. This was very recently answered by Dimitrov, Gao and Habbeger.

Theorem 13. [DGH21] *Let X be a smooth projective geometrically integral curve of genus $g \geq 2$ defined over a number field K of degree d. Let r be the rank of the Mordell-Weil group $J(K)$ of the Jacobian J. Then, there is a constant C depending only on d and g (and not on X), such that*

$$\#X(K) \leq C^{1+r}.$$

Key to this proof is an inequality involving height functions!

Acknowledgements

I would like to thank Borys Kadets for his comments and suggestions on earlier drafts of notes for this lecture series.

References

[DGH21] Vesselin Dimitrov, Ziyang Gao, and Philipp Habegger, *Uniformity in Mordell-Lang for curves*, Ann. of Math. (2) **194** (2021), no. 1, 237–298, DOI 10.4007/annals.2021.194.1.4. MR4276287 ↑8

[Fal83] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), no. 3, 349–366, DOI 10.1007/BF01388432 (German). MR718935 ↑7

[Maz86] Barry Mazur, *Arithmetic on curves*, Bull. Amer. Math. Soc. (N.S.) **14** (1986), no. 2, 207–259, DOI 10.1090/S0273-0979-1986-15430-3. MR828821 ↑7

[Ser89] Jean-Pierre Serre, *Lectures on the Mordell-Weil theorem*, Aspects of Mathematics, E15, Friedr. Vieweg & Sohn, Braunschweig, 1989. Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt. MR1002324 ↑4

[Sil09] Joseph H. Silverman, *The arithmetic of elliptic curves*, 2nd ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009. MR2514094 ↑1, 5

[Sil94] ———, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 151, Springer-Verlag, New York, 1994. MR1312368 ↑4