STAGE Vojta Notes

Niven Achenjang

Fall 2021

These are notes on "Vojta's approach to the Mordell Conjecture", written for my Fall 2021 STAGE talks. They reflect my understanding (or lack thereof) of the material, so are far from perfect. They are likely to contain some typos and/or mistakes, but ideally none serious enough to distract from the mathematics. With that said, enjoy and happy mathing.

To be clear, these are intentionally more detailed than the talks themselves

Contents

Introduction

We aim to give an overview of Bombieri's simplification of Vojta's proof of the Mordell conjecture. In order to keep the whole discussion reasonably brief, we will necessarily omit many details. To fill in the gaps we leave, one may consult [BG06, Chapter 11] and/or [HS00, Part E].

To be specific, we aim to describe a proof of the following statement:

Theorem (Mordell Conjecture)**.** *Let K be a number field, and let C/K be a geometrically irreducible smooth projective curve of genus $g \geq 2$. Then, $\#C(K) < \infty$.*

We begin with a hint of the idea behind the argument. It is harmless to assume on the outset that we have some rational point $P_0 \in C(K)$, and so obtain an embedding $j : C \hookrightarrow J := \mathrm{Pic}^0_{C/K}$ of C into its Jacobian, where $j(P) = [P - P_0]$. The Jacobian is an abelian variety, and so – after choosing some ample, even line bundle $\mathscr{L}$ on J – has a quadratic form $\widehat{h}_{J,\mathscr{L}} : J(\overline{K}) \to \mathbb{R}$ defined on its $\overline{K}$-points. Thus, $J(K) \otimes \mathbb{R} \cong \mathbb{R}^{\operatorname{rank} J(K)}$ becomes a Euclidean space. In particular, the quadratic form $\widehat{h}_{\mathscr{L}}$ induces an inner product $\langle -, - \rangle$ and norm $|\cdot|$ on $J(K) \otimes \mathbb{R}$.

In 1965, Mumford [Mum65] showed that the image of $C(K) \overset{j}{\hookrightarrow} J(K) \to J(K) \otimes \mathbb{R}$ is 'sparse' in the sense that given two points $P, Q \in C(K)$, of large height (and small angle between them), with images $p, q \in J(K) \otimes \mathbb{R}$, one must have $|p| \geq 2\,|q|$ or vice versa. This alone is not enough to prove Mordell, but Vojta improved the situation by further showing that

Theorem (Vojta's inequality)**.** *There are constants $\kappa_1 = \kappa_1(C)$ and $\kappa_2 = \kappa_2(g)$ such that, for $P, Q \in C(\overline{K})$, one has*

$$|P| \geq \kappa_1 \;\; \text{and} \;\; |Q| \geq \kappa_2\,|P| \implies \frac{\langle P, Q \rangle}{|P|\,|Q|} \leq \frac{3}{4}.$$

That is, two 'large' points on the curve cannot be too far away from each either without having small angle between them. This stands in contrast to Mumford's result that any two 'large' points with small angle must be far away. Playing these two off of each other, combined with a simple geometric argument, allows one to show that C may only have finitely many points of large height. Since a curve always has only finitely many 'small' points, this gives Mordell's conjecture.

In the first section of these notes, we will give the details of the geometric argument taking us from Vojta's inequality to Mordell's Conjecture. Afterwards, we will prove Mumford's results, and then describe the way by which one gets from there to Vojta's inequality. Throughout, we will need to appeal to various facts about (canonical) heights on Jacobians of curves which we have not yet seen in this seminar. All of the relevant results are collected in Appendix A. I recommend perusing it before continuing with these notes, and then referring back to it as needed while reading them.

1 The Punch line

Before diving into the proof of Vojta's inequality. Let's take a closer look at how it can be used, along with Mumford's Gap Principle, to prove Mordell's conjecture. For the reader's convenience, we restate these results below.

Notation 1.1. Let K be a number field, and let C/K be a genus $g \geq 2$ curve with basepoint $P_0 \in C(K)$. Let $J = \mathrm{Pic}^0_{C/K}$ be its Jacobian, let $j : C \hookrightarrow J, P \mapsto [P - P_0]$ denote the embedding determined by P_0, and let $\Theta = j(C) + \cdots + j(C) \in \mathrm{Div}(J)$ be the Theta divisor. Let $\Theta^- = [-1]^*\Theta$, so $\widehat{h}_{J,\Theta+\Theta^-}$ gives a quadratic form on $J(\overline{K})$. Let $|\cdot|$ denote the associated norm, so $|x|^2 = \widehat{h}_{J,\Theta+\Theta^-}(x)$, and let $\langle -, - \rangle$ denote the associated bilinear form on $J(\overline{K})$, so

$$\langle x, y \rangle = \frac{1}{2}\left(|x+y|^2 - |x|^2 - |y|^2\right).$$

Abuse of Notation 1.2. We will often be sloppy by implicitly identifying points of $C(\overline{K})$ with their images (under j) in $J(\overline{K})$ or even in $J(\overline{K}) \otimes \mathbb{R}$, writing expressions such that $|P|, \langle P, Q \rangle$ instead of $|j(P)|, \langle j(P), j(Q) \rangle$.

Theorem 1.3. *For any $\varepsilon > 0$, there are constants $B = B(C, P_0, \varepsilon)$ and $\kappa = \kappa(g) \geq 1$ such that for any distinct $P, Q \in C(\overline{K})$ satisfying $|P| \geq |Q| > B$ and*

$$\cos\theta(P,Q) := \frac{\langle P, Q\rangle}{|P|\,|Q|} \geq \frac{3}{4} + \varepsilon$$

(i.e. the angle between P, Q is small) one has

- ***(Mumford's Gap Principle)*** $2\,|Q| \leq |P|$.
- ***(Vojta's Inequality)*** $|P| \leq \kappa\,|Q|$

We first remark that if one only wants to prove the Mordell Conjecture, then Vojta's inequality alone suffices. Indeed,

Corollary 1.4. $\#C(K) < \infty$

Proof. Recall that the kernel of the map $J(K) \to J(K) \otimes \mathbb{R}$ is the torsion subgroup of $J(K)$, which is finite, so it suffices to show that the image of $C(K)$ in $J(K) \otimes \mathbb{R}$ is finite.

For any poits $x, y \in J(K) \otimes \mathbb{R}$ the angle $\theta(x, y)$ between them is defined via

$$\cos\theta(x,y) = \frac{\langle x, y\rangle}{|x|\,|y|} \text{ and } 0 \leq \theta(x,y) \leq \pi.$$

For any point x_0 and angle θ_0, we consider the cone

$$\Lambda_{x_0,\theta_0} := \{x \in J(K) \otimes \mathbb{R} : \theta(x, x_0) < \theta_0\},$$

so any two elements of Λ_{x_0,θ_0} differ by an angle at most $2\theta_0$. We'll show that Vojta's inequality implies that Λ_{x_0,θ_0} can only have finitely elements of $C(K)$ if θ_0 is small.

Indeed, suppose θ_0 so small that $\cos(2\theta_0) > \frac{3}{4}$, and fix some $Q \in \Lambda_{x_0,\theta_0} \cap C(K)$ with $|Q| > B$.[1] Then, by Vojta's inequality, every $P \in \Lambda_{x_0,\theta_0} \cap C(K)$ satisfies $|P| \leq B$, $|P| \leq |Q|$ or $|P| \leq \kappa\,|Q|$, so in fact $|P| \leq \kappa\,|Q|$ always. Thus, the points of $\Lambda_{x_0,\theta_0} \cap C(K)$ have bounded height (bounded by $\kappa^2\,|Q|^2 = \kappa^2 \widehat{h}_{J,\Theta+\Theta^-}(Q)$), and so there are only finitely many points of $C(K)$ in each small cone.

Thus, to win, it suffices to observe that $J(K) \otimes \mathbb{R}$ can be covered by finitely many cones of the form Λ_{x_0,θ_0} with θ_0 small. The quickest way to see this is the observe that such cones are open and a collection of them covers $J(K) \otimes \mathbb{R}$ iff it covers the unit sphere $S = \{|x| = 1\} \subset J(K) \otimes \mathbb{R}$, but this unit sphere is compact. Hence, finitely many such cones will cover it and so cover all of $J(K) \otimes \mathbb{R}$. ■

That's great, but we'd ultimately like to have a uniform bound on the number of points (of large height) of $C(K)$. The argument does not even give us an explicit bound for a single curve C. To remedy this, one uses Mumford's Gap Principle as well as the following geometric fact.

Proposition 1.5. *Fix some $\theta_0 \in (0, \pi/2)$. There is a set $\{x_1, \ldots, x_n\} \subset \mathbb{R}^d$ such that*

$$\mathbb{R}^d = \bigcup_{i=1}^{n} \Lambda_{x_i,\theta_0} \text{ where } \Lambda_{x_i,\theta_0} := \left\{y \in \mathbb{R}^d : \theta(x,y) \leq \theta_0\right\},$$

[1] If this is not possible, the points of C in this cone have bounded height, so there are only finitely many of them.

and furthermore

$$n \leq (1 + \csc(\theta_0/2))^d - 1.$$

Proof. Assume wlog that $|x_i| = 1$ for all i. Let $S \subset \mathbb{R}^d$ denote the unit sphere, and note that the Λ_{x_i,θ_0}'s cover $\mathbb{R}^d$ iff they cover S. If $x \in \Lambda_{x_i,\theta_0}$ and $|x| = 1$, then, by the law of sines (applied to the triangle with vertices $0, x, x_i$), one gets

$$\frac{\sin\theta(x, x_i)}{|x - x_i|} = \frac{1}{|x - 0|}\sin\left(\frac{\pi - \theta(x, x_i)}{2}\right) = \cos\left(\frac{\theta(x, x_i)}{2}\right)$$

which implies that

$$|x - x_i| = \frac{\sin\theta(x, x_i)}{\cos(\theta(x, x_i)/2)} = 2\sin(\theta(x, x_i)/2) \leq 2\sin(\theta_0/2)$$

(above, we used $\sin(2t) = 2\sin(t)\cos(t)$ with $t = \theta(x, x_i)/2$). Thus, letting $r := 2\sin(\theta_0/2)$, and letting $S_{x_i,r} \subset \mathbb{R}^d$ denote the closed ball of radius r centered at x_i, we have

$$S \cap S_{x_i,r} = S \cap \Lambda_{x_i,\theta_0}.$$

Thus, the Λ_{x_i,θ_0}'s cover $\mathbb{R}^d$ iff the $S_{x_i,r}$'s cover $S = S_{0,1}$. Say $x_1, \ldots, x_n$ is a maximal collection of points on the sphere satisfying

$$|x_i - x_j| \geq r \text{ for all } i \neq j.$$

Then, the $S_{x_i,r}$ cover S by maximality, so these give a collection of points for which the Λ_{x_i,θ_0}'s cover $\mathbb{R}^d$. To bound the number of points, note that the $S_{x_i,r/2}$'s are disjoint with union

$$\bigcup_{i=1}^{n} S_{x_i,r/2} \subset S_{0,1+r/2} - S_{0,r/2}$$

by the triangle inequality. Letting $v = \mathrm{vol}(S) = \mathrm{vol}(S_{0,1})$, we then get

$$nv(r/2)^d = \mathrm{vol}\left(\bigsqcup_i S_{x_i,r/2}\right) \leq \mathrm{vol}(S_{0,1+r/2}) - \mathrm{vol}(S_{0,r/2}) = v(1 + r/2)^d - v(r/2)^d,$$

whence the claimed inequality

$$n \leq \left(\frac{1 + r/2}{r/2}\right)^d - 1 = (1 + 2/r)^d - 1 = \left(1 + \frac{1}{\sin(\theta_0/2)}\right)^d - 1. \qquad \blacksquare$$

Exercise. Using Proposition 1.5 and Mumford's Gap Principle, adapt the proof of Corollary 1.4 to obtain the following stronger result:

With B, κ as in Theorem 1.3, if $\Gamma \leq J(\overline{K})$ is finitely generated, then

$$\#\left\{P \in C(\overline{K}) : P \in \Gamma \text{ and } |P| \geq B\right\} \leq (\log_2 \kappa + 1)7^{\mathrm{rank}\,\Gamma}.$$

In particular, taking $\Gamma = J(K)$ provides an explicit upper bound for the number of large points on a given curve.

To extend the above to a uniform bound on the number of large points on a curve of genus $g \geq 2$, one would need to keep closer track of how the constants (especially $B = B(C, P_0, \varepsilon)$) depend on the choice of curve. We will not do this, but if one were to do this, then they could prove a result like

Theorem 1.6 ([Gao21], Theorem 3.2)**.** *Let S be an irreducible variety, and let $\pi : \mathcal{A} \to S$ be an abelian scheme of relative dimension $g \geq 1$. Let $\mathscr{L}$ be a relatively ample line bundle on $\mathcal{A}/S$ such that $[-1]^*\mathscr{L} \simeq \mathscr{L}$. Let $\widehat{h}_{\mathscr{L}} : \mathcal{A}(\overline{\mathbb{Q}}) \to \mathbb{R}_{\geq 0}$ denote the fiber-wise Néron-Tate height. Moreover, let $\mathcal{M}$ be an ample line bundle over a compactification $\overline{S}$ of S, and so obtain a height function $h_{\overline{S},\mathcal{M}} : \overline{S}(\overline{\mathbb{Q}}) \to \mathbb{R}$ downstairs.*

Let $\mathcal{C} \subset \mathcal{A}$ be an irreducible closed subvariety dominating S and suppose that $\mathcal{C} \to S$ is a flat family of curves of genus ≥ 2. Then, there is a constant $c = c(\pi, \mathscr{L}, \mathcal{M} : \mathcal{C}) \geq 1$ so that for any $s \in S(\overline{\mathbb{Q}})$ and subgroup $\Gamma \leq \mathcal{A}_s(\overline{\mathbb{Q}})$ of finite rank $\rho \geq 0$, one has

$$\#\left\{P \in \mathcal{C}_s(\overline{\mathbb{Q}}) \cap \Gamma : \widehat{h}_{\mathscr{L}}(P) > c \max\left\{1, h_{\overline{S},\mathcal{M}}(s)\right\}\right\} \leq c^{\rho}.$$

In these notes, our main goal is Corollary 1.4, and so we will devote the remaining sections to giving an overview of the proofs of Mumford's Gap Principle and Vojta's Inequality.

2 Mumford's Gap Principle

We keep the conventions from Notation 1.1. In particular, C/K is a genus $g \geq 2$ curve over a number field, and it is equipped with a basepoint $P_0 \in C(K)$. We let J denote its Jacobian with embedding $j : C \hookrightarrow J$, and we let $\Theta \in \mathrm{Div}(J)$ be the so-called theta-divisor. Furthermore, $|\cdot|, \langle -, - \rangle$ denote the norm and inner product on $J(\overline{K})$ induced by the ample, symmetric divisor $\Theta + \Theta^-$.

We now recall the goal of this section

Theorem 2.1. *For any $\varepsilon > 0$, there is a constant $B = B(C, P_0, \varepsilon)$ such that for any distinct $P, Q \in C(\overline{K})$ satisfying $|P| \geq |Q| > B$,*

$$\cos\theta(P, Q) \geq \frac{3}{4} + \varepsilon \implies 2\,|Q| \leq |P|\,.$$

The main lemma used in proving this result is sometimes called **Mumford's formula** (Corollary 2.5). For later use in the proof of Vojta's inequality, we will here prove a slight generalization of Mumford's formula.

Mumford's original formula gives an upper bound for the height $h_{C\times C,\Delta}(P, Q)$ of a point on the product $C \times C$ with respect to the diagonal divisor $\Delta \subset C \times C$. Vojta obtained an analogous result for so-called 'Vojta divisors', combinations of the Diagonal along with horizontal and vertical copies of C.

Definition 2.2. Let $C_1 = \mathrm{pr}_1^*(P_0) = P_0 \times C \in \mathrm{Div}(C \times C)$ and let $C_2 = \mathrm{pr}_2^*(P_0) = C \times P_0 \in \mathrm{Div}(C \times C)$. Furthermore, let $\Delta' := \Delta - C_1 - C_2$. A **Vojta divisor** is any divisor of the form

$$V(d_1, d_2, d) := d_1 C_1 + d_2 C_2 + d\Delta' \in \mathrm{Div}(C \times C),$$

with $d_1, d_2, d \in \mathbb{Z}$. ◇

In order to bound the height of points w.r.t. such a divisor, we will need to recall (e.g. from the appendix), the following facts relating various divisors on Jacobians.

Recall 2.3.

(a) Let $h_\Theta : J \xrightarrow{\sim} \widehat{J}$ be the isomorphism induced by the theta divisor, and let $\delta := (\mathrm{id} \times h_\Theta)^* \wp_J \in \mathrm{Div}(J \times J)$ be the "Poincaré bundle viewed on $J \times J$." Then,

$$\delta \sim m^*\Theta - \mathrm{pr}_1^*\Theta - \mathrm{pr}_2^*\Theta \in \mathrm{Div}(J \times J),$$

where $m : J \times J \to J$ is the addition map, and $\mathrm{pr}_1, \mathrm{pr}_2 : J \times J \rightrightarrows J$ are the projection maps. This is from Proposition A.18.

(b) Furthermore, given $x, y \in J(\overline{K})$, one has

$$\widehat{h}_{J\times J,\delta}(x,y) = \widehat{h}_{J\times \widehat{J},\wp_J}(x, h_\Theta(y)) = 2\langle x, y\rangle_\Theta = \langle x, y\rangle$$

via Proposition A.7 combined with Proposition A.18.

Exercise. Instead of going through the lengthy argument for this using the appendix, directly compute this by simply expanding out $\widehat{h}_{J\times J,\delta}(x,y)$ (this also avoids the need for part **(a)** of this recall). For the sake of this exercise, define $\delta := m^*\Theta - \mathrm{pr}_1^*\Theta - \mathrm{pr}_2^*\Theta$.

(c) Finally, $g[P_0] \sim j^*(\Theta^-) \in \mathrm{Div}(C)$ by Corollary A.13. $\odot$

Lemma 2.4. *Let $P, Q \in C(\overline{K})$, and fix a Vojta divisor $V = V(d_1, d_2, d) = d_1C_1 + d_2C_2 + d\Delta'$. Then,*

$$h_{C\times C,V}(P,Q) = \frac{d_1}{2g}|P|^2 + \frac{d_2}{2g}|Q|^2 - d\langle P,Q\rangle + d_1O(|P|) + d_2O(|Q|) + (d_1+d_2+d+1)O(1).$$

Proof. First recall that Remark A.19 shows that $(j\times j)^*\delta \simeq \mathscr{O}_{C\times C}(-\Delta')$. This combined with Recall 2.3**(b)** shows that

$$h_{C\times C,\Delta'}(P,Q) = \widehat{h}_{J\times J,-\delta}(P,Q) + O(1) = -\langle P,Q\rangle + O(1).$$

Next note that

$$h_{C\times C,C_1}(P,Q) = h_{C\times C,\mathrm{pr}_1^*[P_0]}(P,Q) = h_{C,[P_0]}(P) + O(1).$$

By Recall 2.3**(c)**, we know that $g[P_0] \sim j^*(\Theta^-)$, so we must have

$$h_{C,[P_0]}(P) = \frac{1}{g}h_{C,g[P_0]}(P) + O(1) = \frac{1}{g}\widehat{h}_{J,\Theta^-}(P) + O(1)$$

above. Finally, note that $\widehat{h}_{J,\Theta^-} = \frac{1}{2}\left[\widehat{h}_{J,\Theta+\Theta^-} - \widehat{h}_{J,\Theta-\Theta^-}\right]$, and so conclude that

$$h_{C\times C,C_1}(P,Q) = \frac{1}{g}\widehat{h}_{J,\Theta^-}(P) + O(1) = \frac{1}{2g}|P|^2 - \frac{1}{2g}\widehat{h}_{J,\Theta-\Theta^-}(P) + O(1).$$

One similarly shows that

$$h_{C\times C,C_2}(P,Q) = \frac{1}{g}\widehat{h}_{J,\Theta^-}(Q) + O(1) = \frac{1}{2g}|Q|^2 - \frac{1}{2g}\widehat{h}_{J,\Theta-\Theta^-}(Q) + O(1).$$

Thus,

$$h_{C\times C,V}(P,Q)=\frac{d_1}{2g}\left|P\right|^2+\frac{d_2}{2g}\left|Q\right|^2-d\left\langle P,Q\right\rangle-\frac{d_1}{2g}\widehat{h}_{J,\Theta-\Theta^-}(P)-\frac{d_2}{2g}\widehat{h}_{J,\Theta-\Theta^-}(Q)+(d_1+d_2+d+1)O(1).$$

To conclude, we use Proposition A.6 to see that $\widehat{h}_{J,\Theta-\Theta^-}(P)=O(|P|)$ from whence we get the claim. ∎

Corollary 2.5 (**Mumford's Formula**). *For $\Delta=V(1,1,1)$, one has*

$$h_{C\times C,\Delta}(P,Q)=\frac{1}{2g}\left|P\right|^2+\frac{1}{2g}\left|Q\right|^2-\left\langle P,Q\right\rangle+O(|P|+|Q|+1).$$

Corollary 2.6 (**Mumford's Gap Principle**). *Fix some $\varepsilon>0$. There is a constant $B=B(C,P_0,\varepsilon)>0$ such that for any distinct $P,Q\in C(\overline{K})$, satisfying $|P|\geq|Q|>B$, one has*

$$\cos\theta(P,Q)=\frac{\left\langle P,Q\right\rangle}{\left|P\right|\left|Q\right|}\geq\frac{3}{4}+\varepsilon\implies 2\left|Q\right|\leq\left|P\right|.$$

Proof. By Mumford's formula (+ $|P|\geq|Q|>B$ to simplify the big O's a bit), we know

$$\left\langle P,Q\right\rangle+h_{C\times C,\Delta}(P,Q)=\frac{1}{2g}\left|P\right|^2+\frac{1}{2g}\left|Q\right|^2+O(|P|).$$

Since $(P,Q)\notin\operatorname{supp}\Delta$, we may assume that $h_\Delta(P,Q)\geq 0$. Thus, assuming $\cos\theta(P,Q)\geq 3/4+\varepsilon$,

$$\frac{1}{2g}\left|P\right|^2+\frac{1}{2g}\left|Q\right|^2+O(|P|)\geq\left\langle P,Q\right\rangle\geq\left(\frac{3}{4}+\varepsilon\right)\left|P\right|\left|Q\right|.$$

That is,

$$\frac{3}{4}+\varepsilon\leq\frac{1}{2g}\left(\frac{|P|}{|Q|}+\frac{|Q|}{|P|}\right)+O\left(\frac{1}{|Q|}\right)\leq\frac{1}{2g}\left(\frac{|P|}{|Q|}+1\right)+O\left(\frac{1}{|Q|}\right).$$

If $|Q|\gg 0$, the big-O term above will be $<\varepsilon$, and so we will have

$$\frac{3}{4}\leq\frac{1}{2g}\left(\frac{|P|}{|Q|}+1\right).$$

This rearranges to

$$\frac{|P|}{|Q|}\geq\frac{6g}{4}-1\geq\frac{12}{4}-1=2,$$

the desired conclusion. ∎

3 Vojta's Inequality

For the sake of brevity, this section will be missing many details (which can found in [BG06, Chapter 11] or [HS00, Part E]).[2]

We aim to sketch a proof of

[2]This whole section could use some cleaning. If you find a mistake or something that can be improved, let me know.

Theorem 3.1. *There are constants $\kappa_1 = \kappa_1(C)$ and $\kappa_2 = \kappa_2(g)$ such that, for any $P, Q \in C(\overline{K})$ with $|P| \geq |Q| > \kappa_1$, one has*

$$\cos\theta(P,Q) \geq \frac{3}{4} + \varepsilon \implies |P| \leq \kappa_2 |Q|.$$

At a high level, the proof of this inequality is very similar to the proof of Mumford's Gap Principle. Lemma 2.4 gives an equation of the form

$$h_{C\times C,V}(P,Q) = (\text{quadratic in } |P|, |Q|) + (\text{linear in } |P|, |Q|) + O(1)$$

for any fixed Vojta divisor $V = V(d_1, d_2, d)$. One wants to choose d_1, d_2, d carefully so as to force the quadratic term to be very negative. At the same time, if $(P,Q) \notin \operatorname{supp} V$, then $h_{C\times C,V}(P,Q) \geq 0$, so for such P, Q, the linear term must be larger (in absolute value) than the quadratic. This forces a bound on such $|P|, |Q|$.

Part of the difficulty with extending this argument to all P, Q is that the choice of $V = V(d_1, d_2, d)$ used in the end will depend on the points P, Q. Hence, we cannot guarantee that $P, Q \notin \operatorname{supp} V$. In order to combat this, one must do a careful analysis of the height function $h_{C\times C,V}$ in order to produce general (negative) lower bounds on $h_{C\times C,V}(P,Q)$ which still suffice to get a non-trivial height bound in the end. To obtain this, one does two things: **(1)** produces a precise lower bound for $h_{C\times C,V}(P,Q)$ in terms of given local equations cutting out V (i.e. in terms of a section of $\mathscr{O}_{C\times C}(V)$), and **(2)** produces a section giving as small (in magnitude) a lower bound as possible.

For the below sections, keep the conventions from Notation 1.1.

3.1 Vojta Divisors, and lot's of notation

Recall 3.2. A *Vojta divisor* on $C \times C$ is one of the form

$$V(d_1, d_2, d) = d_1 C_1 + d_2 C_2 + dC. \qquad \odot$$

In our attempt to obtain bounds for $h_{C\times C,V}$, we will need to describe things as explicitly as possible. Hence, our first goal is to write V as a difference of two very ample divisors on $C \times C$ in order to relate $h_{C\times C,V}$ directly to the absolute logarithmic height on projective space. This will involve introducing lots of new notation, which is summarized at the end of this section in Table 1.

Fix some $N \geq 2g+1$, so that $N[P_0]$ is very ample on C by Riemann-Roch. Let

$$\varphi_{N[P_0]} : C \hookrightarrow \mathbb{P}^n_K$$

denoted the corresponding closed immersion. This gives rise to a closed embedding

$$\psi = \varphi_{N[P_0]} \times \varphi_{N[P_0]} : C \times C \longrightarrow \mathbb{P}^n \times \mathbb{P}^n$$

so that

$$\psi^* \mathscr{O}_{\mathbb{P}^n\times\mathbb{P}^n}(\delta_1, \delta_2) \simeq \mathscr{O}_{C\times C}(\delta_1 N C_1 + \delta_2 N C_2). \tag{3.1}$$

Corollary 3.3. *In particular, the above equality shows that $\delta_1 N C_1 + \delta_2 N C_2$ is very ample for any*

$\delta_1, \delta_2 \geq 1$. *Furthermore, if $M \gg 0$, then*

$$B := M(C_1 + C_2) - \Delta'$$

is very ample on $C \times C$ as well.

Fix a choice of M as in the above corollary, and let $\varphi_B : C \times C \hookrightarrow \mathbb{P}^m$ be the corresponding closed embedding.

Notation 3.4. We denote the homogeneous coordinates on $\mathbb{P}^m$ (the target of φ_B) by $\vec{y}$ (i.e. by $y_0, y_1, \ldots, y_m$), and we denote those on $\mathbb{P}^n \times \mathbb{P}^n$ (the target of ψ) by $(\vec{x}, \vec{x}')$.

Throughout this argument, we will generally assume that our Vojta divisors $V = V(d_1, d_2, d)$ satisfy the following three axioms

(V1)

$$\delta_i := \frac{d_i + Md}{N} \in \mathbb{Z}_{\geq 1} \text{ for } i = 1, 2.$$

This gives our decomposition

$$V = d_1 C_1 + d_2 C_2 + d\Delta' = N(\delta_1 C_1 + \delta_2 C_2) - dB$$

of V into a difference of two very ample divisors. Note that this can be rephrased as

$$\mathscr{O}_{C\times C}(V) \simeq \psi^* \mathscr{O}_{\mathbb{P}^n \times \mathbb{P}^n}(\delta_1, \delta_2) \otimes \left(\varphi_B^* \mathscr{O}_{\mathbb{P}^m}(d)\right)^{-1}.$$

(V2) The first cohomology groups of the twisted ideal sheaves $\mathscr{I}_{\mathbb{P}^n \times \mathbb{P}^n / \psi(C\times C)}(\delta_1, \delta_2), \mathscr{I}_{\mathbb{P}^m / \varphi_B(C \times C)}(d)$ vanish. This will hold as long as we take $\delta_1, \delta_2, d \gg 0$, and then the natural restriction maps

$$\psi^* : \Gamma\left(\mathbb{P}^n_K \times \mathbb{P}^n_K, \mathscr{O}(\delta_1, \delta_2)\right) \longrightarrow \Gamma(C \times C, \mathscr{O}(\delta_1 N C_1 + \delta_2 N C_2)) \text{ and } \varphi_B^* : \Gamma(\mathbb{P}^m_K, \mathscr{O}(d)) \longrightarrow \Gamma(C \times C, \mathscr{O}(dB))$$

will be surjective.

(V3) $d_1 + d_2 > 4g - 4$ and $d_1 d_2 - g d^2 > \gamma d_1 d_2$ for some $\gamma > 0$.

The significance of this axiom won't be seen until section 3.3.

Not every intermediate result will require all three of these, but you can go ahead and assume them all if you want.

Within these axioms, we have obtained our decompostion of V into a difference of two very ample divisors. Let's see how to use this to explicitly describe sections of $\mathscr{O}_{C\times C}(V)$ using our coordinates $(\vec{x}, \vec{x}')$ and $\vec{y}$ on $\mathbb{P}^n \times \mathbb{P}^n$ and $\mathbb{P}^m$.

Lemma 3.5. *Let V be a Vojta divisor satisfying **(V1)** and **(V2)**. For any global section $s \in \Gamma(C \times C, \mathscr{O}(V))$, there are bihomogeneous polynomials $F_i(\vec{x}, \vec{x}')$, $i = 0, \ldots, m$, of bidegree (δ_1, δ_2), such that*

$$s = \left.\frac{F_i(\vec{x}, \vec{x}')}{y_i^d}\right|_{C \times C} \tag{3.2}$$

for $i = 0, \dots, m$.

Conversely, assume that $F_i(\vec{x}, \vec{x}')$, $i = 0, \dots, m$, are bihomogeneous polynomials of bidegree (δ_1, δ_2) satisfying

$$\frac{F_i(\vec{x}, \vec{x}')}{y_i^d} = \frac{F_j(\vec{x}, \vec{x}')}{y_j^d}$$

on $C \times C$ for every i, j. Then, there is a unique global section $s \in \Gamma(C \times C, \mathscr{O}(V))$ such that (3.2) holds for all i.

Proof. This is essentially a direct consequence of

$$\mathscr{O}_{C\times C}(V) \simeq \psi^* \mathscr{O}_{\mathbb{P}^n \times \mathbb{P}^n}(\delta_1, \delta_2) \otimes (\varphi_B^* \mathscr{O}_{\mathbb{P}^m}(d))^{-1} .$$

Spelled out, first let s be a global section of $\mathscr{O}_{C\times C}(V)$, so $s \otimes \big(y_i^d|_{C\times C}\big)$ is a global section of $\psi^* \mathscr{O}_{\mathbb{P}^n \times \mathbb{P}^n}(\delta_1, \delta_2)$. By **(V2)**, this means that $s \otimes \big(y_i^d|_{C\times C}\big) = F_i(\vec{x}, \vec{x}')|_{C\times C}$ is the restriction of a bihomogeneous polynomial F_i of bidegree (δ_1, δ_2), from which we obtain the first statement.

For the converse, say $F_i(\vec{x}, \vec{x}')$, $i = 0, \dots, m$, are bihomogeneous polynomials of bidegree (δ_1, δ_2) satisfying

$$\frac{F_i(\vec{x}, \vec{x}')}{y_i^d} = \frac{F_j(\vec{x}, \vec{x}')}{y_j^d}$$

on $C \times C$ for all i, j. Let $s_i \in \Gamma((C \times C)_{y_i}, V)$ be the meromorphic section defined by (3.2), where $(C \times C)_{y_i} = \{y_i \neq 0\}$ is the open locus where $y_i \in \Gamma(C \times C, \varphi_B^* \mathscr{O}(1))$ is non-vanishing. Then, the s_i's glue to given a regular section on

$$\bigcup_{i=0}^{m} (C \times C)_{y_i} = (C \times C) \setminus \{y_i = 0 \text{ for all } i\} = C \times C,$$

i.e. they give a global section $s \in \Gamma(C \times C, V)$. ■

To end, we summarize the notation we have introduced below in Table 1.

Symbol	Description
$V = V(d_1, d_2, d)$	A choice of Vojta divisor $V = d_1 C_1 + d_2 C_2 + d\Delta'$.
$\varphi_B : C \times C \hookrightarrow \mathbb{P}^m$	Embedding associated to the very ample divisor $B = M(C_1 + C_2) - d\Delta'$ on $C \times C$. Note: B (and so M) is fixed.
$\vec{y} = [y_0 : \cdots : y_m]$	Homogeneous coordinates on $\mathbb{P}^m$.
$\psi = \varphi_{N[P_0]} \times \varphi_{N[P_0]} : C \times C \hookrightarrow \mathbb{P}^n \times \mathbb{P}^n$	Embedding associated to the pair $(N[P_0], N[P_0])$ of very ample divisors on C. Note: N is fixed.
$(\vec{x}, \vec{x}')$	Homogeneous coordinates on $\mathbb{P}^n \times \mathbb{P}^n$
δ_1, δ_2	Integers so that $$V = d_1 C_1 + d_2 C_2 + d\Delta' = N(\delta_1 C_1 + \delta_2 C_2) - dB$$ when V satisfies **(V1)**.
$\gamma > 0$	Number so that $d_1 d_2 - g d^2 > \gamma d_1 d_2$, when V satisfies **(V3)**.

Table 1: Notation used throughout the proof of Vojta's inequality

3.2 A lower bound for $h_{C\times C,V}$

In the previous section, we obtained a description for sections of $\mathscr{O}_{C\times C}(V)$ in terms of coordinates $\vec{y}$ on $\mathbb{P}^m$ and $(\vec{x},\vec{x}')$ on $\mathbb{P}^m\times\mathbb{P}^m$. Let us not use this description to product an explicit lower bound for $h_{C\times C,V}$ at any point $(P,Q)\in C\times C$.

Recall 3.6. Let X/K be a smooth K-variety, and let $x\in X$ be a closed point. Let $\kappa(x):=\mathscr{O}_{X,x}/\mathfrak{m}_x$ denote its residue field. The **tangent space** $T_{X,x}$ of X at x is the $\kappa(x)$-vector space

$$T_{X,x}:=\mathrm{Der}_{\kappa(x)}(\mathscr{O}_{X,x},\kappa(x))\simeq\mathrm{Hom}_{\kappa(x)}(\Omega_{\mathscr{O}_{X,x}/\kappa(x)},\kappa(x))$$

consisting of all $\kappa(x)$-derivations $\partial:\mathscr{O}_{X,x}\to\kappa(x)$. An element of $T_{X,x}$ is called a tangent vector. $\odot$

Consider two points $P,Q\in C(\overline{K})$. Let ∂,∂' be non-zero tangent vectors at P,Q. Abbreviate

$$\partial_i:=\frac{1}{i!}\partial^i\ \text{ and }\ \partial'_i:=\frac{1}{i!}\partial'^i.$$

Any differential operator on $\mathscr{O}_{C\times C,(P,Q)}$ of degree k with values in $\kappa(P,Q)$ is a homogeneous polynomial of degree k in the variables ∂,∂' with coefficients in K.[3] In fact, since $C\times C$ is irreducible (so $K(C\times C)=\mathrm{Frac}\,\mathscr{O}_{C\times C,(P,Q)}$), all such operators acts on $K(C\times C)$.

Remark 3.7. If $f_1,\dots,f_r$ are rational functions on C, then Leibniz's rule tells us that

$$\partial_i(f_1\dots f_r)=\sum_{i_1+\cdots+i_r=i}(\partial_{i_1}f_1)\dots(\partial_{i_r}f_r). \qquad \circ$$

Definition 3.8. Let $s\in\Gamma(C\times C,\mathscr{O}(V))$ be a nonzero global section. A pair $(i_1^*,i_2^*)\in\mathbb{N}^2$ is called **admissible for** (P,Q) iff

$$\partial_{i_1^*}\partial'_{i_2^*}s(P,Q)\neq 0,$$

but $\partial_{i_1}\partial'_{i_2}s(P,Q)=0$ for all $i_1\le i_1^*$ and $i_2\le i_2^*$ with $(i_1,i_2)\neq(i_1^*,i_2^*)$. That is, (i_1^*,i_2^*) give the 'first nonzero Taylor coefficient' of s near (P,Q), so admissibility is a measure of the order of vanishing of s at (P,Q).

In order to make sense of this, one should choose a trivialization of $\mathscr{O}_{C\times C}(V)$ near (P,Q). Note however that admissibility is independent of the choice of trivialization and of the choice of ∂,∂' (different choices only differ by a scalar since $\dim T_{C,P}=1=\dim T_{C,Q}$). $\diamond$

Recall that $(\vec{x},\vec{x}')$ denote the homogeneous coordinates on $\mathbb{P}^n\times\mathbb{P}^n$. Hence,

$$\xi_{ij}:=\left(\frac{x_i}{x_j}\right)\bigg|_C\ \text{ and }\ \xi'_{ij}:=\left(\frac{x'_i}{x'_j}\right)\bigg|_C$$

gives well-defined non-zero rational functions on C, for $i,j=0,\dots,n$.

Notation 3.9. We write

$$\vec{\xi}_j:=(\xi_{0j},\xi_{1j},\dots,\xi_{nj})=\left(\frac{x_0}{x_j},\dots,\frac{x_n}{x_j}\right)$$

[3] Really to make proper sense of this, you need ∂,∂' to be differential operators in neighborhoods of the points P,Q, and not just at these points. I didn't wanna bother being careful about this technical point in these notes.

for the vector with components $\xi_{ij} = x_i/x_j$ for $i = 0, \dots, n$. We similarly define $\vec{\xi}'_j$.

Now, note that, with respect to the choices made in section 3.1 (in particular axiom **(V1)**), we can choose an explicit height function associated to V. First choose a finite extension L/K so that $P, Q \in C(L)$. Recall that

$$\mathscr{O}_{C\times C}(V) \simeq \psi^* \mathscr{O}_{\mathbb{P}^n \times \mathbb{P}^n}(\delta_1, \delta_2) \otimes (\varphi_B^* \mathscr{O}_{\mathbb{P}^m}(d))^{-1} .$$

We set

$$\begin{aligned} h_{C\times C,V}(P,Q) &:= \sum_{v \in M_L} \max_{\substack{H,H' \\ |H|=\delta_1, |H'|=\delta_2}} \min_{\substack{I \\ |I|=d}} \log \left| \frac{\vec{x}^H \vec{x}'^{H'}}{\vec{y}^I}(P,Q) \right|_v \\ &= \sum_{v \in M_L} \max_{j,j'} \min_i \log \left| \frac{x_j^{\delta_1} x_{j'}^{\delta_2}}{y_i^d}(P,Q) \right|_v , \end{aligned}$$

where M_L denote the set of all places of L.

Notation 3.10. For $s \in \mathscr{O}_{C\times C}(V)$, Lemma 3.5 gives bihomogeneous polynomials $F_i(\vec{x}, \vec{x}')$ of bidegree (δ_1, δ_2) with

$$s = \frac{F_i(\vec{x}, \vec{x}')}{y_i^d} \text{ on } C \times C \text{ for } i = 0, \dots, m.$$

We let $h(\vec{F})$ denote the height of the projective point whose coordinates are given by all the coefficients of $F_0, \dots, F_m$.

For $v \in M_K$, we let j_v be the index j for which $|\xi_{j0}(P)|_v$ is largest, and similarly

$$j'_v := \arg\max_j \left| \xi'_{j0}(Q) \right|_v .$$

Lemma 3.11. *Let s be a nonzero global section of $\mathscr{O}_{C\times C}(V)$, and let (i_1^*, i_2^*) be admissible for s at (P,Q). With the notation introduced above, one has*

$$\begin{aligned} h_V(P,Q) \geq &- h(\vec{F}) - n \log((\delta_1 + n)(\delta_2 + n)) \\ &- \sum_{v \in M_L} \max_{\{i_\lambda\}} \left(\sum_\lambda \max_\nu \log |\partial_{i_\lambda} \xi_{\nu j_v}(P)|_v \right) \\ &- \sum_{v \in M_L} \max_{\{i'_\lambda\}} \left(\sum_\lambda \max_\nu \log \left| \partial'_{i'_\lambda} \xi'_{\nu j'_v}(Q) \right|_v \right) \\ &- (\delta_1 + \delta_2 + i_1^* + i_2^*) \log 2 \end{aligned}$$

where $\{i_\lambda\}$ and $\{i'_\lambda\}$ run over all partitions of i_1^* and i_2^*.

Proof. We fix trivializations of $\mathscr{O}_{\mathbb{P}^n}(1)$ at P, Q and of $\mathscr{O}_{\mathbb{P}^m}(1)$ at (P,Q). These give rise to trivializations of all line bundles in question, and in particular, of $\mathscr{O}_{C\times C}(V)$.

Recall that

$$h_V(P,Q) = -\sum_v \max_i \min_{j,j'} \log \left| \frac{y_i^d}{x_j^{\delta_1} x'^{\delta_2}_{j'}}(P,Q) \right|_v .$$

The only i, j, j' that matter above are those for which $x_j(P), x'_{j'}(Q), y_i(P,Q) \neq 0$, so assume this is the case. Admissibility + Leibniz give (the second equality in)

$$\partial_{i_1^*}\partial'_{i_2^*}F_i(\vec{\xi}_j, \vec{\xi}'_{j'}) = \partial_{i_1^*}\partial'_{i_2^*}\left(\frac{y_i^d}{x_j^{\delta_1}x_{j'}'^{\delta_2}}s\right)(P,Q) = \left(\frac{y_i^d}{x_j^{\delta_1}x_{j'}'^{\delta_2}}\partial_{i_1^*}\partial'_{i_2^*}s\right)(P,Q)$$

(all other terms vanish). Since $(\partial_{i_1^*}\partial'_{i_2^*}s)(P,Q) \neq 0$, this gives

$$\begin{aligned} h_V(P,Q) &= 0 - \sum_v \max_i \min_{j,j'} \log\left|\frac{y_i^d}{x_j^{\delta_1}x_{j'}'^{\delta_2}}(P,Q)\right|_v \\ &= -\sum_v \log\left|(\partial_{i_1^*}\partial'_{i_2^*}s)(P,Q)\right|_v - \sum_v \max_i \min_{j,j'} \log\left|\frac{y_i^d}{x_j^{\delta_1}x_{j'}'^{\delta_2}}(P,Q)\right|_v \quad \text{by the product formula} \\ &= -\sum_v \max_i \min_{j,j'} \log\left|\left(\frac{y_i^d}{x_j^{\delta_1}x_{j'}'^{\delta_2}}\partial_{i_1^*}\partial'_{i_2^*}s\right)(P,Q)\right|_v \\ &= -\sum_v \max_i \min_{j,j'} \log\left|\partial_{i_1^*}\partial'_{i_2^*}F_i(\vec{\xi}_j, \vec{\xi}'_{j'})(P,Q)\right|_v. \end{aligned}$$

Now, the number of monomials of F_i is bounded by $\binom{\delta_1+n}{n}\binom{\delta_2+n}{n} \leq (\delta_1+n)^n(\delta_2+n)^n$, so

$$\begin{aligned} h_V(P,Q) \geq &- h(\vec{F}) - n\log((\delta_1+n)(\delta_2+n)) \\ &- \sum_v \min_j \max_{|\vec{l}|=\delta_1} \log\left|\partial_{i_1^*}\vec{\xi}_j^{\vec{l}}(P)\right|_v \\ &- \sum_v \min_{j'} \max_{|\vec{l}'|=\delta_2} \log\left|\partial'_{i_2^*}\vec{\xi}_{j'}'^{\vec{l}'}(Q)\right|_v \end{aligned}$$

(get this by splitting F_i into its monomials and bounding each of them).

For each v, we take the minimum with respect to j, j'. We may instead take $j = j_v$ and $j' = j'_v$. Consider $\log\left|\partial_{i_1^*}\vec{\xi}_j^{\vec{l}}(P)\right|_v$ for $v \in M_L$. By Leibniz,

$$\partial_{i_1^*}\vec{\xi}_j^{\vec{l}} = \sum \prod_{\nu=0}^{n}\prod_{\mu=1}^{\ell_\nu} \partial_{i_{\mu\nu}}\xi_{\nu j} \;\text{ where }\; \sum_{\mu\nu} i_{\mu\nu} = i_1^*.$$

The total number of pairs $\mu\nu$ above is $\delta_1 = \left|\vec{l}\right|$, so stars and bars tells us that there are $\binom{\delta_1+i_1^*-1}{i_1^*} \leq 2^{\delta_1+i_1^*}$ possibilities for $i_{\mu\nu}$. We are interested in the case $j = j_v$. Since $|\xi_{j_v 0}(P)|_v$ is the largest $|\xi_{j0}(P)|_v$, we have

$$|\xi_{\nu j_v}(P)|_v = \left|\frac{\xi_{\nu 0}(P)}{\xi_{j_v 0}(P)}\right|_v \leq 1 \;\text{ for all }\; \nu.$$

Thus, terms with $i_{\mu\nu} = 0$ will contribute $\leq \log 1 = 0$, and so we can ignore them in obtaining the bound[4]

$$\log \left|\partial_{i_1^*} \vec{\xi}^{\,l}_{j_v}(P)\right|_v \leq \max_{\{i_\lambda\}} \left(\sum_\lambda \max_\nu \log \left|\partial_{i_\lambda} \xi_{\nu j_v}(P)\right|_v \right) + \varepsilon_v(\delta_1 + i_1^*) \log 2,$$

where $\varepsilon_v = [L_v : \mathbb{R}]/[L : \mathbb{Q}]$ if $v \mid \infty$ and $\varepsilon_v = 0$ otherwise, and where $\{i_\lambda\}$ runs over all partitions of i_1^*. An analogous estimate holds for the sum involving $\vec{\xi}^{\,l}_{j'}$, and this suffices to get the claim. ■

The sums appearing in the above lemma are a little inconvenient to work with. In order to simplify life, one can apply a theorem of Eisenstein [BG06, Theorem 11.4.1] bounding Taylor coefficients of algebraic functions in order to arrive at

Lemma 3.12 ([BG06], Lemma 11.6.7)**.** *There exists a finite subset $Z \subset C(\overline{K})$ such that for $P \notin Z$, one has*

$$\sum_{v \in M_L} \max_{\{i_\lambda\}} \left(\sum_\lambda \max_\nu \log \left|\partial_{i_\lambda} \xi_{\nu j_v}(P)\right|_v \right) = O\left(i_1^* \left|P\right|^2 + i_1^* \right),$$

with the max running over all partitions $\{i_\lambda\}$ of i_1^*. The implied constant is independent of P and i_1^*.

Corollary 3.13. *Fix some $(P, Q) \in C \times C$ with $P, Q \notin Z$. Let s be a nonzero global section of $\mathscr{O}_{C \times C}(V)$, and let (i_1^*, i_2^*) be admissible for s at (P, Q). Then,*

$$h_{C \times C, V}(P, Q) \geq -h(\vec{F}) - O\left(i_1^* \left|P\right|^2 + i_2^* \left|Q\right|^2 + i_1^* + i_2^* \right) - O(\delta_1 + \delta_2).$$

3.3 A section of $\mathscr{O}_{C \times C}(V)$ of small height

Corollary 3.13 shows us that if we want a useful lower bound for $h_{C \times C, V}$, then we'll want a section of $\mathscr{O}_{C \times C}(V)$ of small height. To produce such a section, we will reduce the question of constructing sections to a problem about integral solutions of linear transformations, and then apply a lemma due to Siegel on producing small solutions to integral linear equations.

Namely, we will apply the following

Theorem 3.14 (**Siegel's Lemma**, 1929)**.** *Let a_{ij}, $i = 1, \ldots, M$ and $j = 1, \ldots, N$ be rational integers, not all 0, bounded by B and suppose $N > M$. Then, the homogeneous linear system*

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1N}x_N & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2N}x_N & = & 0 \\ \vdots & & \vdots & & \vdots & & & & \vdots \\ a_{M1}x_1 & + & a_{M2}x_2 & + & \ldots & + & a_{MN}x_N & = & 0 \end{array}$$

has a solution $x_1, \ldots, x_N \in \mathbb{Z}$, not all 0, bounded by

$$\max_i |x_i| \leq \left\lfloor (NB)^{\frac{M}{N-M}} \right\rfloor.$$

[4] $\max\{x_i\} \leq \sum x_i \leq n \max\{x_i\}$ (n terms) so taking logs gives

$$\max\{\log x_i\} \leq \log\left(\sum x_i\right) \leq \log n + \max\{\log x_i\}.$$

Proof. Let $A = (a_{ij})$. We may assume no row is identically 0. For a positive integer k, let

$$T_k := \left\{\vec{x} \in \mathbb{Z}^N : 0 \leq x_i \leq k \text{ for all } 1 \leq i \leq N\right\}.$$

Let S_m^+ denote the sum of the positive entries in the mth row of A, and let S_m^- denote the sum of the negative entries. For $\vec{x} \in T_k$ and $\vec{y} := A\vec{x}$, we have

$$kS_m^- \leq y_m \leq kS_m^+.$$

Let

$$T_k' := \left\{\vec{y} \in \mathbb{Z}^M : kS_m^- \leq y_m \leq kS_m^+ \text{ for all } 1 \leq m \leq M\right\}.$$

If $B_m := \max_n |a_{mn}|$ is the largest absolute value in the mth row, then $S_m^+ - S_m^- \leq NB_m$, so T' has at most $\prod_m (NkB_m + 1)$ elements. Now, choose k so that $\#T > \#T'$, i.e.

$$\prod_m (NkB_m + 1) < (k+1)^N$$

(note $N > M$), e.g. let $k = \left\lfloor \prod_m (NB_m)^{1/(N-M)} \right\rfloor$ and use $NkB_m + 1 < NB_m(k+1)$ to see this choice of k works.. By pigeonhole, we then get two different points $\vec{x}', \vec{x}'' \in T$ with $A\vec{x}' = A\vec{x}''$, and $\vec{x} := \vec{x}' - \vec{x}''$ is a solution in integers with

$$\max_n |x_n| \leq k \leq \left\lfloor \prod_m (NB)^{1/(N-M)} \right\rfloor = \left\lfloor (NB)^{M/N-M} \right\rfloor. \qquad \blacksquare$$

Corollary 3.15. *Let K be a number field of degree d contained in $\mathbb{C}$ with $|\cdot|$ the usual absolute value on $\mathbb{C}$. Let $M, N \in \mathbb{N}$ with $0 < M < N$. There are positive constants C_1, C_2 such that for any nonzero $M \times N$ matrix A with entries $a_{mn} \in \mathscr{O}_K$, there is some $\vec{x} \in \mathscr{O}_K^N \setminus \{\vec{0}\}$ with $A\vec{x} = 0$, and*

$$H(\vec{x}) \leq C_1 (C_2 NB)^{\frac{M}{N-M}},$$

where $B := \sup_{\sigma,m,n} |\sigma(a_{mn})|$ and σ ranging over embeddings $K \hookrightarrow \mathbb{C}$.

We do not prove the corollary here, but the basic idea is to use that $\mathscr{O}_K \cong \mathbb{Z}^d$ in order to expand things out to a situation with $\mathbb{Z}$-coefficients where you can apply the form of Siegel's lemma dealing with rational integral matrices.

Now, let's see how to use this to produce a small section. We first estimate the sizes of some relevant cohomology groups.

Lemma 3.16 ([HS00], Lemma E.6.1)**.** *We have*

$$\dim \Gamma(C \times C, \psi^* \mathscr{O}(\delta_1, \delta_2)) = (N\delta_1 + 1 - g)(N\delta_2 + 1 - g),$$

and, for $d_1 + d_2 > 4g - 4$,

$$\dim \Gamma(C \times C, \mathscr{O}(V)) \geq d_1 d_2 - gd^2 + O(d_1 + d_2).$$

(The second estimate above is the reason we introduced axiom **(V3)**)

One proves this by applying Riemann-Roch for surfaces, using that d_1, d_2 are sufficiently large to show that the Serre dual of the relevant line bundle has no global sections (since it will have negative intersection with an ample divisor). To get a strict equality in the first case above, one needs to know the H^1 term vanishes; this comes from an application of Kodaira vanishing.

Once one has computed the dimensions of these cohomology groups, they are in position to apply Siegel's lemma.

Lemma 3.17. *There are two positive constants C_4, C_5 independent of d_1, d_2, d and γ with the following property. Let V be a Vojta divisor satisfying **(V1)**,**(V2)**,**(V3)**, and $d_1, d_2 \geq C_4/\gamma$. Then, there is a nonzero global section s of $\mathscr{O}_{C\times C}(V)$ such that the polynomials $F_0, \dots, F_m$ in Lemma 3.5 may be chosen with*

$$h(\vec{F}) \leq C_5 \frac{d_1 + d_2}{\gamma}.$$

Proof Sketch. We want to apply Siegel's lemma to get a section of small height, and so we'll need to transfer the equations in Lemma 3.5 into a linear system of equation with coefficients in K.

We consider C as a curve in $\mathbb{P}^N_K$ of degree N (via the closed embedding $\varphi_{N[P_0]}$), and we may also assume, by a linear change of coordinates, that the projection $p(\vec{x}) = (x_0 : x_1 : x_2)$ maps C birationally onto a curve in $\mathbb{P}^2_K$. This reduces the number of linear equations to be considered in the application of Siegel's lemma. Moreover, we may also assume that $p(C)$ is explicitly given by a homogeneous polynomial

$$f(x_0, x_1, x_2) = a_0 + a_2 x_2 + \cdots + a_{N-1} x_2^{N-1} + x_2^N$$

with $a_i \in K[x_0, x_1]$ homogeneous of degree $N - i$.

The point of this simplification is that the monomials in $F_i \in K[x_0, x_1, x_2; x_0', x_1', x_2']$ with x_2- and x_2'-degrees $< N$ are linearly independent. This is important because to apply Siegel, we'll need to be able to estimate the dimensions of the spaces of unknowns and of solutions of the linear system described by the equations $F_i/y_i^d = F_j/y_j^d$. Now, one writes $y_i = p_i(\vec{x}; \vec{x}')$ with $p_i \in K[x_0, x_1, x_2; x_0', x_1', x_2']$ (with $\deg_{x_2} p_i < N$ and $\deg_{x_2'} p_i < N$) and uses this to obtain a linear system (in the coefficients of the polynomials F_i) whose solutions give sections of $\mathscr{O}_{C\times C}(V)$. If one carefully keeps track of the sizes of the coefficients appearing in this system, and uses Lemma 3.16 to estimate the dimension of the spaces of solutions and the number of unknowns, then they will obtain the claimed result. ■

3.4 "Roth's Lemma," and the Proof of Vojta's inequality

There is one last technical result needed before one can prove Vojta's inequality. The previous section allows us to obtain a section $s \in \Gamma(C \times C, \mathscr{O}(V))$ with small height, but staring at Corollary 3.13 shows that this is not enough; we also need s to vanish to low degree at (P, Q), so it has a small admissible pair. To guarantee this, one uses a lemma due to Roth in order to strengthen 3.17 and so obtain

Lemma 3.18 ([BG06], Lemma 11.8.6)**.** *There is a constant $C_6 > 0$, independent of d_1, d_2, d and γ such that for $0 < \varepsilon < 1/\sqrt{2}$, for any Vojta divisor satisfying **(V1)**,**(V2)**,**(V3)** with*

$$\gamma d_2 \geq C_4 \text{ and } d_2 \leq \varepsilon^2 d_1$$

and for any $P, Q \in C(\overline{K})$ with

$$\min\left(d_1 h_{N[P_0]}(P), d_2 h_{N[P_0]}(Q)\right) \geq C_6 \frac{d_1}{\gamma \varepsilon^2}$$

there is a nonzero global section s of $\mathscr{O}_{C\times C}(V)$ with an admissible pair (i_1^*, i_2^*) at (P, Q) such that

$$h(\vec{F}) \leq C_5 \frac{d_1 + d_2}{\gamma} \quad \textit{and} \quad \frac{i_1^*}{d_1} + \frac{i_2^*}{d_2} \leq 4N\varepsilon.$$

With this last ingredient taken for granted, we may conclude:

Theorem 3.19 (Vojta's inequality)**.** *There are constants $\kappa_1 = \kappa_1(C)$ and $\kappa_2 = \kappa_2(g) > 1$ such that, for $P, Q \in C(\overline{K})$, one has*

$$|P| \geq \kappa_1 \ \textit{ and } \ |Q| \geq \kappa_2 |P| \implies \frac{\langle P, Q\rangle}{|P|\,|Q|} \leq \frac{3}{4}.$$

Proof. I'm gonna be sloppy with some of the big-O stuff because being careful is not my forté.

Note that the set Z to be avoided in Corollary 3.13 is finite, so $P, Q \notin Z$ if $|P|, |Q| \gg 0$. Fix a small positive $\gamma_0 < 1$ and some $D \in \mathbb{N}$. Let $V = V(d_1, d_2, d)$ be a Vojta divisor with

$$d_1 = \frac{D}{|P|^2}\sqrt{g+\gamma_0} + O(1), \ \ d_2 = \frac{D}{|Q|^2}\sqrt{g+\gamma_0} + O(1), \ \text{ and } \ d = \frac{D}{|P|\,|Q|} + O(1).$$

The $O(1)$'s above are to insure that $d_1, d_2, d, \delta_1, \delta_2$ are all nonzero natural numbers. Note that this V satisfies **(V1)** (because of the $O(1)$'s), **(V2)** (by choosing $D \gg 0$), and **(V3)** as

$$d_1 d_2 - g d^2 \geq \gamma d_1 d_2 \ \text{ for } \ \gamma = \frac{\gamma_0}{g+\gamma_0} + o(1)$$

(with the $o(1)$ tending to 0 as $D \to \infty$).

Now, lemma 2.4 gives

$$\begin{aligned} h_{C\times C, V}(P, Q) &= \frac{d_1}{2g}|P|^2 + \frac{d_2}{2g}|Q|^2 - d\langle P, Q\rangle + O\left(d_1|P| + d_2|Q| + d_1 + d_2\right) \\ &= D\left(\frac{\sqrt{g+\gamma_0}}{g} - \frac{\langle P, Q\rangle}{|P|\,|Q|}\right) + O\left(\frac{D}{|P|} + \frac{D}{|Q|}\right). \end{aligned}$$

Corollary 3.13 combined with Lemma 3.17 then gives

$$-O\left(\frac{d_1+d_2}{\gamma} + i_1^*|P|^2 + i_2^*|Q|^2 + i_1^* + i_2^* + \delta_1 + \delta_2\right) \leq h_{C\times C, V}(P, Q) = D\left(\frac{\sqrt{g+\gamma_0}}{g} - \frac{\langle P, Q\rangle}{|P|\,|Q|}\right) + O\left(\frac{D}{|P|} + \frac{D}{|Q|}\right).$$

We can further manipulate this:

$$\begin{aligned} \frac{\langle P, Q\rangle}{|P|\,|Q|} &\leq \frac{\sqrt{g+\gamma_0}}{g} + O\left(\frac{1}{|P|} + \frac{1}{|Q|}\right) + \frac{1}{D}O\left(\frac{d_1+d_2}{\gamma} + i_1^*|P|^2 + i_2^*|Q|^2 + i_1^* + i_2^* + \delta_1 + \delta_2\right) \\ &= \frac{\sqrt{g+\gamma_0}}{g} + O\left(\frac{1}{|P|} + \frac{1}{|Q|} + \frac{1}{|P|^2} + \frac{1}{|Q|^2} + \frac{i_1^*}{d_1} + \frac{i_2^*}{d_2} + \frac{i_1^* + i_2^*}{D}\right) \\ &= \frac{\sqrt{g+\gamma_0}}{g} + O\left(\frac{1}{|P|} + \frac{1}{|Q|} + \frac{i_1^*}{d_1} + \frac{i_2^*}{d_2} + \frac{i_1^* + i_2^*}{D}\right). \end{aligned}$$

Since $|Q| \geq |P|$ in the end, this further simplifies to

$$\frac{\langle P, Q\rangle}{|P|\,|Q|} \leq \frac{\sqrt{g+\gamma_0}}{g} + O\left(\frac{1}{|P|}\right) + \left(\frac{i_1^*}{d_1} + \frac{i_2^*}{d_2}\right) O\left(1 + \frac{1}{|P|^2}\right). \tag{3.3}$$

Now, to apply Lemma 3.18, we'd like some small ε so that

$$\frac{d_2}{d_1} \leq \varepsilon^2 \text{ and } \min\left(d_1 h_{N[P_0]}(P), d_2 h_{N[P_0]}(Q)\right) \geq C_6 \frac{d_1}{\gamma\varepsilon^2}. \tag{3.4}$$

The first of these translates to

$$\frac{|P|}{|Q|} \leq \varepsilon + o(1). \tag{3.5}$$

For the second, recall from the proof of Lemma 2.4 that $h_{N[P_0]}(P) = \frac{N}{2g}|P|^2 + O(|P|) + O(1)$ (and similarly for Q), so

$$d_1 h_{N[P_0]}(P), d_2 h_{N[P_0]}(Q) \geq \frac{DN}{2g}\sqrt{g+\gamma_0}$$

while

$$C_6 \frac{d_1}{\gamma\varepsilon^2} = \frac{1}{\varepsilon^2} O\left(\frac{D}{|P|^2}\right).$$

If we first fix $\varepsilon < 1/\sqrt{2}$ satisfying (3.5), then we get (the second part of) (3.4) by simply taking $|P|$ large, say $|P| \geq \kappa_1$. This puts us in a position to apply Lemma 3.18 in order to obtain

$$\frac{i_1^*}{d_1} + \frac{i_2^*}{d_2} \leq 4N\varepsilon$$

which implies (recall (3.3))

$$\frac{\langle P, Q\rangle}{|P|\,|Q|} \leq \frac{\sqrt{g+\gamma_0}}{g} + O\left(\frac{1}{|P|}\right) + 4N\varepsilon O\left(1 + \frac{1}{|P|^2}\right) \leq \frac{\sqrt{g+\gamma_0}}{g} + O\left(\frac{1}{\kappa_1}\right) + O(\varepsilon)$$

(when $|P| \leq \varepsilon\,|Q|$, and so we'll take $\kappa_2 := 1/\varepsilon$). Since $g \geq 2$, $\sqrt{g}/g \leq \sqrt{2}/2 < 3/4$, so if we take γ_0, ε small enough and κ_1 large enough, the above will say $\cos\theta(P, Q) \leq 3/4$. ■

Appendices

A Canonical Heights and Jacobians

We first briefly recall the definition of the Picard variety of a smooth variety X.

Definition A.1. Let K be a field, and let X/K be a variety. Two line bundles $\mathscr{L}_1, \mathscr{L}_2 \in \operatorname{Pic}(X)$ are said to be **algebraically equivalent** if there is an irreducible smooth variety T and a line bundle $\mathscr{L}$ on $X \times T$ so that

$$\mathscr{L}_1 \cong \mathscr{L}|_{X_{t_1}} \text{ and } \mathscr{L}_2 \cong \mathscr{L}|_{X_{t_2}},$$

for some $t_1, t_2 \in T(K)$. We will denote algebraic equivalence by $\mathscr{L}_1 \equiv \mathscr{L}_2$. We let

$$\operatorname{Pic}^0(X) := \{\mathscr{L} \in \operatorname{Pic}(X) : \mathscr{L} \equiv \mathscr{O}_X\}$$

denote the group of line bundles algebraically equivalent to the trivial bundle. $\diamond$

Fact. Let K be a field, and let X/K be an irreducible, smooth projective variety. To keep things simple, assume we have a base point $P_0 \in X(K)$. Then, the functor $\operatorname{Pic}^0_{X/K}$ defined by

$$\operatorname{Pic}^0_{X/K}(T) := \left\{ (\mathscr{L} \in \operatorname{Pic}(X \times T), \iota) \;\middle|\; \begin{array}{c} \mathscr{L}_t \in \operatorname{Pic}^0(X_{\kappa(t)}) \text{ for any } t \in T \\ \iota : \mathscr{L}_{P_0} \xrightarrow{\sim} \mathscr{O}_T \end{array} \right\} \cong \frac{\operatorname{Pic}^0(X \times T)}{\operatorname{pr}_2^* \operatorname{Pic}(T)}$$

is representable by a scheme, also denoted by $\operatorname{Pic}^0_{X/K}$. Elements of $\operatorname{Pic}^0_{X/K}(T)$ are called **subfamilies of** $\operatorname{Pic}^0(X)$ **parameterized by** T.

Since this functor is representable there is, in particular, a universal line bundle, the **Poincaré bundle** $\wp$ on $X \times \operatorname{Pic}^0_{X/K}$ such that $\wp$ is a subfamily of $\operatorname{Pic}^0(X)$ parameterized by $\operatorname{Pic}^0_{X/K}$ and for any $\mathscr{L} \in \operatorname{Pic}^0_{X/K}(T)$, there is a unique morphism $\varphi_{\mathscr{L}} : T \to \operatorname{Pic}^0_{X/K}$ with $(\mathrm{id}_X \times \varphi)^* (\wp) = \mathscr{L}$. This $\wp \in \operatorname{Pic}^0_{X/K}(\operatorname{Pic}^0_{X/K})$ is the family corresponding to the identity morphism $\operatorname{Pic}^0_{X/K} \xrightarrow{=} \operatorname{Pic}^0_{X/K}$.

See [BG06, 8.4.6 and Theorem 8.4.13] for a discussion of the above fact, as well as references for its proof.

Applying the above fact with $X = C$ a curve yields its Jacobian $J := \operatorname{Pic}^0_{C/K}$. Applying it with $X = A$ an abelian variety, yields the dual abelian variety $\widehat{A} := \operatorname{Pic}^0_{A/K}$.

Remark A.2 (Jacobians). Note that if $X = C$ is a curve, then the above fact yields the Jacobian $J := \operatorname{Pic}^0_{C/K}$ of C. Furthermore, the basepoint $P_0 \in C(K)$ yields a natural family $P \mapsto \mathscr{O}_C(P - P_0)$ of degree 0 line bundles on C, parameterized by C. This is the family giving rise to the usual Abel-Jacobi map $j : C \to J$. $\circ$

Remark A.3 (Dual abelian varieties). If $X = A$ is an abelian variety, then $\widehat{A} := \operatorname{Pic}^0_{A/K}$ is its dual abelian variety (and $P_0 = 0 \in A(K)$). Given any line bundle $\mathscr{L} \in \operatorname{Pic}(A)$, one gets a morphism $\varphi_{\mathscr{L}} : A \to \widehat{A}$ corresponding to the family $A \ni x \longmapsto \tau_x^* \mathscr{L} \otimes \mathscr{L}^{-1}$, where $\tau_x : A \to A, a \mapsto a + x$ is the translation by x map. This morphism is surjective iff $\mathscr{L}$ is ample. $\circ$

A.1 Heights

For this section, we largely follow [BG06, Chapter 9], especially sections 9.1 – 9.3.

Fix a number field K, and an abelian variety A/K.

Definition A.4. For any $n \in \mathbb{Z}$, let $[n] : A \to A$ denote the multiplication-by-n morphism. We call a line bundle $\mathscr{L} \in \operatorname{Pic}(A)$ **even** if $[-1]^*\mathscr{L} \simeq \mathscr{L}$ and **odd** if $[-1]^*\mathscr{L} \simeq \mathscr{L}^{-1}$. ⋄

Fact. $\mathscr{L}$ above is odd $\iff \mathscr{L} \in \operatorname{Pic}^0(A)$.

Recall A.5 (Canonical/Néron-Tate Height Machinery). There is a homomorphism

$$\begin{array}{rccc} \widehat{h}: & \operatorname{Pic}(A) & \longrightarrow & \mathbb{R}^{A(\overline{K})} \\ & \mathscr{L} & \longmapsto & \widehat{h}_{A,\mathscr{L}} \end{array}$$

assigning to each line bundle on A a height function on its $\overline{K}$-points so that

(a) $\widehat{h}_{A,\mathscr{L}}$ is the unique quadratic function satisfying $\widehat{h}_{A,\mathscr{L}} = h_{A,\mathscr{L}} + O(1)$ and $\widehat{h}_{A,\mathscr{L}}(0) = 0$.

(b) if $\varphi : A \to B$ is a morphism between abelian varieties, then

$$\widehat{h}_{A,\varphi^*(\mathscr{L})} = \widehat{h}_{B,\mathscr{L}} \circ \varphi - \widehat{h}_{B,\mathscr{L}}(\varphi(0))$$

for all $\mathscr{L} \in \operatorname{Pic}(B)$.

(c) If $\mathscr{L}$ is ample, then for any $D, B > 0$, we have

$$\#\left\{x \in A(\overline{K}) \mid \widehat{h}_{A,\mathscr{L}}(x) < D \text{ and } [\kappa(x) : K] < B\right\} < \infty.$$

(d) If $\mathscr{L} \in \operatorname{Pic}(A)$ is odd, then $\widehat{h}_{A,\mathscr{L}} : A(\overline{K}) \to \mathbb{R}$ is a linear form, i.e. a homomorphism.

(e) If $\mathscr{L} \in \operatorname{Pic}(A)$ is even, then $\widehat{h}_{A,\mathscr{L}} : A(\overline{K}) \to \mathbb{R}$ is a quadratic form.

(f) If $\mathscr{L}$ is even and ample, then $\widehat{h}_{A,\mathscr{L}} \geq 0$. Furthermore, $\widehat{h}_{A,\mathscr{L}}(P) = 0 \iff P$ is torsion. ⊙

When working over an abelian variety, we can obtain well-defined heights without having to worry about bounded functions.

Now, say $\mathscr{L} \in \operatorname{Pic}(A)$. Then $\widehat{h}_{A,\mathscr{L}}$ is a quadratic function with associated symmetric bilinear form

$$\langle x, y\rangle_{\mathscr{L}} := \frac{1}{2}\left(\widehat{h}_{A,\mathscr{L}}(x+y) - \widehat{h}_{A,\mathscr{L}}(x) - \widehat{h}_{A,\mathscr{L}}(y)\right).$$

If $\mathscr{L}$ is furthermore even and ample, then $\widehat{h}_{A,\mathscr{L}}$ is a quadratic form, and it also has an associated norm

$$|x|_{\mathscr{L}} := \sqrt{\langle x, x\rangle_{\mathscr{L}}} = \sqrt{\widehat{h}_{A,\mathscr{L}}(x)}.$$

These both extend naturally to $A(K) \otimes \mathbb{R} \cong \mathbb{R}^{\operatorname{rank} A(K)}$, giving it the structure of a Euclidean space. Note that, by Recall A.5**(f)**, this inner product on $A(K) \otimes \mathbb{R}$ is positive definite.

We will use this language to relate heights for odd line bundles to those for even line bundles. We will then show that, in fact, all bilinear forms as above come from the height $\widehat{h}_{A\times\widehat{A},\wp}$ on $A \times \widehat{A}$ associated to the Poincaré bundle $\wp$.

Proposition A.6. *Let $\mathscr{L}' \in \operatorname{Pic}^0(A)$ be an odd line bundle, and let $\mathscr{L} \in \operatorname{Pic}(A)$ be an even, ample line bundle. Then,*

$$\widehat{h}_{A,\mathscr{L}'} = O\left(|\cdot|_{\mathscr{L}}\right) = O\left(\widehat{h}^{1/2}_{A,\mathscr{L}}\right).$$

Proof. Since $\mathscr{L}$ is ample, the map $\varphi_{\mathscr{L}} : A \to \widehat{A}$ is surjective, so we can find some $a \in A(\overline{K})$ so that $\mathscr{L}' = \varphi_{\mathscr{L}}(a) = \tau_a^*\mathscr{L} \otimes \mathscr{L}^{-1} \in \operatorname{Pic}(A_{\overline{K}}) = \widehat{A}(\overline{K})$. Thus,

$$\widehat{h}_{A,\mathscr{L}'}(x) = \widehat{h}_{A,\tau_a^*\mathscr{L}\otimes\mathscr{L}^{-1}}(x) = \widehat{h}_{A,\mathscr{L}}(x+a) - \widehat{h}_{A,\mathscr{L}}(0+a) - \widehat{h}_{A,\mathscr{L}}(x) = \langle x, a\rangle_{\mathscr{L}} \leq 2\,|x|_{\mathscr{L}}\,|a|_{\mathscr{L}} = 2\,|a|_{\mathscr{L}}\sqrt{\widehat{h}_{A,\mathscr{L}}(x)}$$

where we've applied both Recall A.5**(b)** and Cauchy-Schwarz above. This yields the claim. ∎

Proposition A.7. *Let $\mathscr{L} \in \operatorname{Pic}(A)$ with symmetric bilinear form $\langle -,-\rangle_{\mathscr{L}}$. Let $\wp \in \operatorname{Pic}(A \times \widehat{A})$ be the Poincaré class of A, and let $\varphi_{\mathscr{L}} : A \to \widehat{A}$ be the associated polarization. Then,*

$$2\langle a, a'\rangle_{\mathscr{L}} = \widehat{h}_{A\times\widehat{A},\wp}(a, \varphi_{\mathscr{L}}(a')) \quad \textit{for all } a, a' \in A(\overline{K}).$$

Proof. This is a single chain of equalities

$$\begin{aligned}
2\langle a, a'\rangle_{\mathscr{L}} &= \widehat{h}_{A,\mathscr{L}}(a+a') - \widehat{h}_{A,\mathscr{L}}(a') - \widehat{h}_{A,\mathscr{L}}(a) && \text{by definition}\\
&= \widehat{h}_{A,\varphi_{\mathscr{L}}(a')}(a) && \text{by the reasoning in the proof of Proposition A.6}\\
&= \widehat{h}_{A\times\widehat{A},\wp}(a, \varphi_{\mathscr{L}}(a')) - \widehat{h}_{A\times\widehat{A},\wp}(0, \varphi_{\mathscr{L}}(a')) && \text{since } \wp|_{A\times\{\varphi_{\mathscr{L}}(a')\}} \simeq \varphi_{\mathscr{L}}(a')\\
&= \widehat{h}_{A\times\widehat{A},\wp}(a, \varphi_{\mathscr{L}}(a')) - \widehat{h}_{\widehat{A},\wp|_{0\times\widehat{A}}}(\varphi_{\mathscr{L}}(a'))\\
&= \widehat{h}_{A\times\widehat{A},\wp}(a, \varphi_{\mathscr{L}}(a')) && \text{since } \wp|_{0\times\widehat{A}} = 0 \in \operatorname{Pic}(\widehat{A}) \text{ by definition of } \operatorname{Pic}^0_{A/K}(\widehat{A}).
\end{aligned}$$

∎

Remark A.8. To get rid of the annoying factor of two in the above proposition statement, one could replace $2\langle a, a'\rangle_{\mathscr{L}}$ with $\langle a, a'\rangle_{\mathscr{L}+[-1]^*\mathscr{L}}$. ○

A.2 Jacobians

For this section, we largely follow [BG06, Chapter 8], especially section 8.10.

We would like to specialize the above discussion of heights, to the case where $A = \operatorname{Jac}(C)$ is the jacobian of a curve.

Setup A.9. Let K be a field, and let C/K be an irreducible smooth projective curve of genus $g \geq 1$. Fix a basepoint $P_0 \in C(K)$, let $J = \operatorname{Pic}^0_{C/K}$ be its Jacobian, and let $j : C \to J, P \mapsto [P - P_0]$ be the Abel-Jacobi map.

Definition A.10. The **theta divisor** on J is

$$\Theta := \underbrace{j(C) + \cdots + j(C)}_{g-1} \subset J.$$

This is an irreducible, ample divisor such that the associated map

$$\varphi_\Theta : J \xrightarrow{\sim} \widehat{J}$$

from J to its dual abelian variety is an isomorphism. $\diamond$

Notation A.11. We let $\Theta^- := [-1]^*\Theta = -j(C) - \cdots - j(C)$ denote the pullback of Θ by the multiplication by -1 map on J.

Note that $\Theta + \Theta^-$ is an even (ample) divisor while $\Theta - \Theta^-$ is an odd (non-ample) divisor.

One of the benefits of introducing the divisor Θ^- is that it behaves predictably under pullback to C. For example,

Proposition A.12. *For all $(P_1, \dots, P_g) \in C(K)^g$, one has*

$$\sum_{i=1}^{g} [P_i] \sim j_a^*(\Theta^-) \quad \textit{where} \quad j_a(P) := j(P) - a,$$

and $a := [P_1 + \cdots + P_g] - g[P_0] \in J(K)$.

Proof. See [BG06, Proposition 8.10.15] ■

Corollary A.13. *Taking $P_i = P_0$ for $i = 1, \dots, g$ above, we obtain*

$$g[P_0] \sim j^*(\Theta^-).$$

In the context of Jacobians, there are two Poincaré classes: $\wp_C \in \mathrm{Pic}^0(C \times J)$ and $\wp_J \in \mathrm{Pic}^0(J \times \widehat{J})$. We would like to relate these to each other and to the theta divisor Θ. Our main tool for doing this will be the seesaw theorem, stated below

Theorem A.14 (Seesaw Theorem). *Let X be proper and let T be an arbitrary variety. Let $\mathscr{L}$ be a line bundle on $X \times T$. Then,*

(1) $S = \left\{ t \in T : \mathscr{L}|_{X \times \{t\}} \simeq \mathscr{O}_X \textit{ is trivial} \right\}$ *is a closed subvariety of T.*

(2) $\mathscr{L}|_{X \times S} = \mathrm{pr}_2^* \mathscr{M}$ *for some line bundle $\mathscr{M}$ on S.*

Proof. See [Mum08] ■

In other words, if you have a line bundle that is trivial on vertical fibers (fibers above T), then it is really the pullback of some line bundle on the base. Hence, if you want to show two line bundles on a product are one-in-the-same, it can often suffice to show that they agree fiberwise (by applying seesaw to their difference).

Notation A.15. Let $C_1 = \mathrm{pr}_1^*(P_0) = P_0 \times C \in \mathrm{Div}(C \times C)$, let $C_2 = \mathrm{pr}_2^*(P_0) = C \times P_0 \in \mathrm{Div}(C \times C)$, and let $\Delta' := \Delta - C_1 - C_2 \in \mathrm{Div}(C \times C)$.

Proposition A.16. *Let Δ be the diagonal in $C \times C$. Then,*

$$(\mathrm{id}_C \times j)^* (\wp_C) \simeq \mathscr{O}_{C \times C}(\Delta').$$

Proof. First note that, for any $P \in C$,

$$(\mathrm{id}_C \times j)^* (\wp_C)|_{C\times\{P\}} \simeq \mathscr{O}_C(P - P_0).$$

This is a formal consequence of the universality of $\wp_C$. The natural map $j : C \to J = \mathrm{Pic}^0_{C/K}$ picks out a family of degree 0 line bundles on C, and forming the above pullback simply recovers the corresponding family. Similarly, note that, when $P \neq P_0$,

$$\mathscr{O}_{C\times C}(\Delta')|_{C\times\{P\}} = \mathscr{O}_{C\times C}(\Delta - C_1 - C_2)|_{C\times\{P\}} \simeq \mathscr{O}_C(P - P_0).$$

Let $\mathscr{L} = (\mathrm{id}_C \times j)^*(\wp_C) \otimes \mathscr{O}_{C\times C}(-\Delta')$. By Seesaw, Theorem A.14,

$$S := \left\{P \in C : \mathscr{L}_{C\times\{P\}} \text{ is trivial}\right\} \subset C$$

is a closed subvariety. We have just seen that S contains the dense open $C \setminus \{P_0\}$, so we conclude that $S = C$, and hence – again by seesaw – that $\mathscr{L} = \mathrm{pr}_2^* \mathscr{M}$ for some $\mathscr{M} \in \mathrm{Pic}(C)$. To finish, we note that

$$\mathscr{M} \simeq (\mathrm{id}_C \times j)^* (\wp_C)|_{\{P_0\}\times C} \otimes \mathscr{O}_{C\times C}(-\Delta')|_{\{P_0\}\times C} \simeq \mathscr{O}_C$$

is trivial, as $(\mathrm{id}_C \times j)^* (\wp_C)|_{\{P_0\}\times C} \simeq \mathscr{O}_C = \mathscr{O}_C(P_0 - P_0) \simeq \mathscr{O}_{C\times C}(\Delta')|_{\{P_0\}\times C}$. ■

The next two propositions are similarly proved via see-saw arguments. Their proofs are omitted here, but can be found in [BG06, Section 8.10].

Proposition A.17. *Let $m : J \times J \to J$ be addition, let $\mathrm{pr}_1, \mathrm{pr}_2 : J \times J \rightrightarrows J$ be the projection maps; for $\delta := m^*\Theta^- - \mathrm{pr}_1^* \Theta^- - \mathrm{pr}_2^* \Theta^- \in \mathrm{Div}(J \times J)$, we have*

$$\mathscr{O}_{C\times J}\left((j \times \mathrm{id}_J)^* \delta\right) = \wp_C^{-1}.$$

Proposition A.18. *Let $\varphi_\Theta, \varphi_{\Theta^-} : J \longrightarrow \widehat{J}$, and let $\delta := m^*\Theta^- - p_1^*\Theta^- - p_2^*\Theta^- \in \mathrm{Pic}(J \times J)$. Then,*

$$(\mathrm{id}_J \times \varphi_{\Theta^-})^* (\wp_J) = \mathscr{O}_{J\times J}(\delta) = (\mathrm{id}_J \times \varphi_\Theta)^* (\wp_J).$$

Furthermore, $\delta \sim m^*\Theta - \mathrm{pr}_1^* \Theta - \mathrm{pr}_2^* \Theta$.

Remark A.19. Combining Propositions A.16 and A.17, one sees that

$$\mathscr{O}_{C\times C}\left((j \times j)^* \delta\right) = \mathscr{O}_{C\times C}(-\Delta'). \tag{A.1}$$

Because the above fact is actually used in these notes, but Proposition A.17 is not proven here, we will get an alternate, direct proof of it which does not go through the connection to C's Poincaré bundle.

Alternate Proof of (A.1). This is another see-saw argument. We recall from the proof of Proposition A.16 that for any $P \in C$, we have

$$\mathscr{O}_{C\times C}(-\Delta')|_{C\times\{P\}} \simeq \mathscr{O}_C(P_0 - P).$$

The same better be true of $(j \times j)^* \delta$. One sees that, letting $\iota_P : C \hookrightarrow C \times C, Q \mapsto (Q, P)$

$$(j \times j)^* \delta|_{C \times \{P\}} \sim (m \circ (j \times j) \circ \iota_P)^* \Theta^- - (\mathrm{pr}_1 \circ (j \times j) \circ \iota_P)^* \Theta^- - (\mathrm{pr}_2 \circ (j \times j) \circ \iota_P)^* \Theta^-.$$

We now compute these compositions. First,

$$m \circ (j \times j) \circ \iota_P(Q) = m \circ (j \times j)(Q, P) = j(Q) + j(P) \implies m \circ (j \times j) \circ \iota_P = j_{j(P)}.$$

Similarly, one computes $\mathrm{pr}_1 \circ (j \times j) \circ \iota_P = j$ and $\mathrm{pr}_2 \circ (j \times j) \circ \iota_P = j(P)$. Using Proposition A.12 to compute these pullbacks, we see

$$(j \times j)^* \delta|_{C \times \{P\}} \sim (g[P_0] - j(C)) - (g[P_0]) - (0) = -j(C) = [P_0 - P],$$

so, by See-saw, $\mathscr{O}_{C \times C}\left((j \times j)^* \delta\right), \mathscr{O}_{C \times C}(-\Delta')$ must differ by the pullback (along pr_2) of some line bundle on C. To check that this bundle (on C) is trivial, one computes that $\mathscr{O}_{C \times C}\left((j \times j)^* \delta\right), \mathscr{O}_{C \times C}(-\Delta')$ agree also on $\{P_0\} \times C$. ■

○

References

[BG06] Enrico Bombieri and Walter Gubler. *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2006. 1, 7, 14, 16, 19, 21, 22, 23

[Gao21] Ziyang Gao. Recent developments of the uniform mordell-lang conjecture, 2021. 5

[HS00] Marc Hindry and Joseph H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. An introduction. 1, 7, 15

[Mum65] David Mumford. A remark on Mordell's conjecture. *Amer. J. Math.*, 87:1007–1016, 1965. 2

[Mum08] David Mumford. *Abelian varieties*, volume 5 of *Tata Institute of Fundamental Research Studies in Mathematics*. Published for the Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition. 22