Diophantine Problems of Avoiding Unlikely Intersections

Francesco Ballini
Jesus College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Michaelmas 2024

Abstract

Masser and Zannier have proved in 2019 that "most" abelian varieties, defined over $\overline{\mathbb{Q}}$, are not isogenous to any jacobian of a curve. In other terms, "most" points of $\mathcal{A}_g(\overline{\mathbb{Q}})$ "avoid" the countably many "isogenous images" of the Torelli locus.

We generalise their statement, also in the context of powers of modular curves, in a number of ways, including allowing for countable families "to be avoided" rather than the fixed Torelli locus and for a smaller source of "avoiding points", rather than the whole of $\mathcal{A}_g$, to be chosen.

Our second main topic is "double unlikely intersections": we generalise a result by Marché-Maurin related to curves in powers of the multiplicative group and we prove an analogous version in the context of powers of modular curves.

Acknowledgements

First, I must immensely thank my DPhil supervisor Prof. Pila for the many mathematical suggestions and for always being available whenever I asked for his advice. The questions that I have tackled in this document have been, for the most part, pointed out to me by him. During my stay at the University of Oxford I have learned a lot, thanks to his example, about how to do research.

I am in large debt also to Prof. Zannier, who has always supported me during these years and who has been a constant source of inspiring Mathematics. I owe him my gratitude as he gave me the possibility to discover a broad range of topics.

I am grateful to Prof. Hrushovski and Prof. Rössler, who have been the assessors for my Transfer of Status, and to Prof. Daw and Prof. Bays, who assessed my Confirmation of Status, for their generous amount of valuable feedback, both in terms of content and in terms of presentation.

Among the many people who have been fundamental to me during these years at the Mathematical Institute, I definitely must say a "thank you" to my Model Theory peers, Michal, the only mathematician with street credibility, Benedikt, for being so nice, funny and kind, and Leo, a true politician, who have been good friends, great mathematicians to discuss with and able organisers of pub trips on the other side of the road: our organisational power didn't shine as it would have deserved and I will firmly claim it all had been a conspiracy.

I must thank Wojciech, with Delta as well, for always bringing a (usually inappropriate) laugh at the table and Quinn, for his enlightened philosophical thoughts. I must say "thanks" to Clara for the many nights out ranting.

I thank Andrés, my first friend in Oxford, and Almudene, for the parties and the nights spent together. I thank Dario, who has now long left, Filippo, for our meetings in the kitchens of the Department, Ismael, for being an unpredictable party boy, and Marc, who speaks an unbelievable italian.

I must thank my Nightshift friends: Jhon, a real dreamer and unstoppable singer, who completely changed my last year in this town, Michael, great mathematician and even better guitarist, a serious but lovely guy with a passion for annoying his dorm mates at 1am, Miguel, our long lost hardworking drummer, Javi, our number 1 fan and a boy with a huge heart, and Diego, who is extremely kind and nice.

A "thank you" goes to the Rivals: Harry Rocco, our charismatic leader and excellent songwriter, and Jacob, great bassist, friend with an amazing sense of humour and comrade in many

adventures.

A big thanks to my Biology friends: Monica, who has been my partner in drama for all these years, Rency, wanderer of the night, and Marco, for the evenings at the Harcourt Arms and at the Jolly Farmers. Even if we spent only a month together in Oxford, I must thank Yuri, who gifted me with amazing discussions, great friends and even a room for three years.

Another big thanks to the new generation of biologists, Filippo, who has been a crazy source of countless hobbies and deep esoteric evolutionary reasonings, Antonio, who saw with me the best and the worst of this town, Ben, with his animals and his passion for everything, Travis, with his detached way of dissecting and discussing every single aspect of reality, and Reuben, with a kind and warm heart, for the nights at the Maths Department.

I need to thank my live-in landlord Cary, who has been a real source of inspiration with all of his deep theories about everything. I am glad we both have been "things' people" and we could discuss about abstract stuff for many hours so often.

I must thank Alessio, the most paranoid human being who had ever set foot in Oxford, who gave me one of the most unreal years of my life and who has been a great friend, one of the few who could openly tell me off whenever I did something inappropriate.

I am grateful to Oliviero, who is, as of now, the most controversial person I have ever met, for the amount of fun we had together and for giving me his gown.

I need to thank the many people I have spent so much time with during the various conferences, discussing about drama, card games, going out, clubbing in inappropriate places and sometimes even Maths: Laura, for basically being my academic godmother and for getting a generous amount of my worst side, Fabrizio, who knows how to party seriously, Amos, for his imagination, Gabriel Dill, who taught me a lot, Julian Demeio, possibly the person I have spent most nights sleeping in the same room with, Simone Coccia, for being out of touch with reality, Luca, for his passionate enthusiasm, Nicola, for knowing about everything, Davide Lombardo, who has been an infinite source of trivia and Maths, Francesco Maria Saettone, for reasons that cannot be stated publicly, and Gessica, for her prophecies.

I now thank some of my many italian friends here: Pietro, who has left a while ago after creating a community, Alessandro, who once I was told is "the only guy in town who is more italian than you", Carlo, for being a literal angel, Giuliano, who is a lot of fun, turmoil, energy, insults and whatever, Bianca, for being so full of ideas and creativity and yet so caring, Filippo Volpin, for his ""good"" jokes, Filippo De Santis, for being so smart and sensitive, Rosario, an enigma, Stefano, with his dry humour, Vittorio, so mysterious yet so nice, and Nicola, who is

the most erudite person I have met in town.

I thank Stefania, for the nights out talking about our turbulent lives.

I also must thank a newer generation among these people who come and leave: Fabio, who is one of the most sensitive and witty guys I've met, Tamara, the sweetest person in town, Elisabetta, deeply rebel and intelligent, Andrea, whose patience with our endless Maths blabbering has been stoic, Carlo, for our pointless nights out, Mattia, for being a good friend, and Chiara, God knows for what.

I can finally say "thank you" to the people I have left back in Italy.

Prima di tutto, ringrazio la mia famiglia: mia madre, per essersi sempre presa il peggio delle mie frustrazioni, e mio padre, per i suoi saggi e fastidiosi consigli. Ringrazio mio nonno, per avermi sempre tenuto in considerazione.

Ringrazio i miei amici di Castegnato: Togno, per il suo esemplare cinismo, Verze, per la sua bontà e anche per la Play, Mangi, per essere sempre pieno di idee irrealistiche, Giulia, per le sue imitazioni e il suo spirito intraprendente, Raffo, laconico e senza passaporto, Marina, per le sue convinzioni, Trap, umano incomprensibile a cui devo dei soldi, Claudia, dolcissima, Break, per il suo charme, Cannone, scheggia impazzita, Chicco, per i suoi deliri, Misha, per la sua sincerità, e, ovviamente, Nino, mio fratello.

Ringrazio Alessandro Braga, con cui ho riallacciato un'amicizia sepolta per tanto tempo.

Ringrazio i miei amici di Brescia: Fano, per essere sempre lo stesso gnarello di sempre, Picchio, per essere ancora più severamente radicale di un tempo, Tomma, per i tempi passati, Adrian, per il suo occhio attento, Jacques, per la sua cultura e il suo carattere, Lori Corda, per la sua costante allegria e non per le sue battute, Meri, per le sue idee, Rossa, per il suo essere tagliente, Paolo Rassu, per il suo buon cuore, Marta, per le sue mine, Stefania, per la sua carica e per la sua parlata epica, e Giulia Toffoli, per essere sempre la più chill della situa.

Ringrazio Skandar, per tutte le vicende assurde.

Ringrazio i miei amici di Pisa: Dario, perché chi altri, Irene, per essere sempre e comunque presa bene, Toti, dolcissimo e irrequieto, Cecc, sensibile e filosofico, oppure compagnone e poetico, Bere, per le infinite domande sempre interessanti, Roberto, per essere true, e Marco, a cui devo molto.

Ringrazio Antonio, per essere sempre stato un faro per me, e Francesco Sala, per la sua ironia

insuperabile.

Ringrazio Benzo, anche se ultimamente ci siamo visti poco. Ringrazio Fabio Ferri, che mi ha insegnato a cucinare.

Contents

1 Introduction **1**

1.1 Avoiding problems 1

1.2 Double unlikely intersections 4

1.3 Structure of the document 5

2 Preliminaries **6**

2.1 Definable sets 6

2.2 O-minimal structures 10

2.3 A problem of Lang 13

2.4 Basics on elliptic curves 17

2.5 The moduli space of elliptic curves 19

2.6 Basics on heights 24

2.7 Heights and varieties 28

3 Avoiding problems **34**

3.1 The Zilber-Pink Conjecture 34

3.2 Some related conjectures 36

3.3 Avoiding problems and literature 37

3.4 First basic results and proofs 39

3.5 Points avoiding rational lines 52

3.6 Some general principles of avoiding problems 56

3.7 Geometric degrees 59

3.8 Classification of weakly special subvarieties 64

3.9 Avoiding problems in powers of modular curves 73

3.10 Avoiding problems in the moduli space of abelian varieties 89

4 Higher dimensional avoiding problems **104**

4.1 Examples of lines avoiding points 104

4.2 Lines avoiding curves 108

4.3 Curves avoiding curves 118

5 Tangencies in powers of the multiplicative group **124**
5.1 Special subvarieties and weakly special subvarieties . . . 125
5.2 Bounds on the degree . . . 128
5.3 General reductions for weakly special subvarieties . . . 130
5.4 Bounds on the height: functoriality . . . 133
5.5 Bounds on the height: Puiseux series . . . 135
5.6 Final arguments . . . 142

6 Tangencies in powers of modular curves **144**
6.1 Heights: infinite part . . . 148
6.2 Heights: finite part (sketch) . . . 152
6.3 Heights: finite part . . . 156

7 Further results **169**
7.1 Identically isogenous curves . . . 169
7.2 Avoiding problems in abelian varieties . . . 180
7.3 Avoiding problems and Hecke orbits . . . 184

Chapter 1

Introduction

1.1 Avoiding problems

Masser and Zannier have proved in 2019, in [42], that "most" abelian varieties are not isogenous to any jacobian of a curve, in the following sense.

The coarse moduli space of principally polarised abelian varieties of dimension g is itself a variety, denoted with $\mathcal{A}_g$. The *Torelli locus* $\mathcal{T}_g \subseteq \mathcal{A}_g$, accounting for the jacobians of semistable curves of genus g, is a proper subvariety of $\mathcal{A}_g$ whenever $g \geq 4$.
We denote with $\mathcal{T}_g^{(n)}$, for an arbitrary positive integer n, the *isogenous images* of $\mathcal{T}_g$, in the sense that $p \in \mathcal{T}_g^{(n)}$ if and only if there exist $q \in \mathcal{T}_g$ such that p and q represent abelian varieties which are related by an isogeny of degree n.
The result of Masser-Zannier is the assertion that "most" points of $\mathcal{A}_g(\overline{\mathbb{Q}})$ do not belong to the countable union

$$\bigcup_{n\geq 1} \mathcal{T}_g^{(n)}(\overline{\mathbb{Q}}).$$

Let us clarify the meaning of "most". Since our aim is "ordering" the $\overline{\mathbb{Q}}$-points of varieties, we appeal to the Weil height, which measures the "arithmetic complexity" of such points.
Let V be a variety defined over $\overline{\mathbb{Q}}$ and let d be a positive integer. We say that the property (P) holds for "most" points of V of degree $\leq d$ whenever

$$\lim_{T\to+\infty} \frac{|\{p \in V(\overline{\mathbb{Q}}) \mid h(p) < T, [\mathbb{Q}(p) : \mathbb{Q}] \leq d \text{ and } p \text{ satisfies (P)}\}|}{|\{p \in V(\overline{\mathbb{Q}}) \mid h(p) < T, [\mathbb{Q}(p) : \mathbb{Q}] \leq d\}|} = 1$$

where both the numerator and the denominator in the fraction above are natural numbers by the Northcott Property.

It turns out that there is nothing special about the Torelli locus $\mathcal{T}_g$. We quote the result by Masser-Zannier; we say that two points $p, q \in \mathcal{A}_g$ are "isogenous" if they represent isogenous abelian varieties.

Theorem (Masser-Zannier)**.** *Let $\mathcal{H} \subseteq \mathcal{A}_g$ be a proper subvariety. There is a positive integer d_0 such that, for every positive integer $d \geq d_0$, "most" points of $\mathcal{A}_g(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $\mathcal{H}(\overline{\mathbb{Q}})$.*

Remark. *We point out that the same result with $\overline{\mathbb{Q}}$ replaced by $\mathbb{C}$ is essentially immediate, for instance by measure-theoretic reasons.*

The main results of this document are generalisations of this Theorem in a few directions and different settings.

We let $Y(1)$ be the coarse moduli space of elliptic curves; as a quasi-projective variety, such space is isomorphic to the affine space $\mathbb{A}^1$. First, we wish to allow $\mathcal{H}$ to vary in countable families of bounded "algebro-geometric" degree and bounded degree of field of definition; second, we wish to restrict our source of "avoiding points".

We now give a simple statement showcasing these two principles.

Proposition. *Let $\mathcal{F} = \{X_1, X_2, \ldots\}$ be a set of lines of $Y(1)^2 \cong \mathbb{A}^2$ defined over $\mathbb{Q}$. Let $Y \subseteq Y(1)^2$ be a curve such that, for every pair of positive integers (i, n), the condition*

$$Y \not\subseteq X_i^{(n)}$$

holds. Then, there is a positive integer d_0 such that, for every positive integer $d \geq d_0$, "most" points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of

$$\bigcup_{i \geq 1} X_i(\overline{\mathbb{Q}})$$

i.e. any point lying on a rational line.

Remark. *We can observe that all our hypotheses (e.g. a bounded "algebro-geometric" degree, a bounded degree of the field of definition and the condition of Y not being a subvariety of any $X_i^{(n)}$) are essential.*

We quote our most general version for $Y(1)^r$, Theorem 3.6.7.

Theorem. *Let a, b, c, r be positive integers and let $\mathcal{F} = \{X_1, X_2, \ldots\}$ be a set of subvarieties of $Y(1)^r$, each of dimension $\leq a$, "algebro-geometric" degree $\leq b$ and field of definition of degree $\leq c$ over $\mathbb{Q}$.*
Let $Y \subseteq Y(1)^r$ be an irreducible subvariety such that, for every pair of positive integers i and n, we have that $Y \cap X_i^{(n)}$ not Zariski dense in Y.
Then, there is a positive integer d_0 such that, for any $d \geq d_0$, "most" points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $X_i(\overline{\mathbb{Q}})$ for any positive integer i.

The proof is similar to the original Masser-Zannier argument in its use of the Pila-Zannier method: the Pila-Wilkie Theorem 2.2.5 is combined with suitable functional transcendence results in order to provide upper bounds which contradict lower bounds coming from Galois-theoretic reasoning and the isogeny estimates Theorem 3.2.2.
The main new ingredients are arguments required to bound the height of a variety among the X_i (or, more precisely, the height of a point representing such a variety in an appropriate

Chow variety) in terms of familiar quantities and a form of isogeny estimates taking into account weakly special subvarieties, our Proposition 3.8.14; in particular, in the Masser-Zannier argument it is essential for "most" points of $\mathcal{A}_g(\overline{\mathbb{Q}})$ not to be contained in any positive dimensional proper weakly special subvariety and such property cannot hold in $Y(1)^r$.

We can obtain a similar result, with an additional restriction on the variety Y, for the ambient variety $\mathcal{A}_g$. This is our Theorem 3.10.1.

Theorem. *Let a, b, c, r be positive integers and let $\mathcal{F} = \{X_1, X_2, \ldots\}$ be a set of subvarieties of $\mathcal{A}_g$, each of dimension $\leq a$, "algebro-geometric" degree $\leq b$ and field of definition of degree $\leq c$ over $\mathbb{Q}$.*
Let $Y \subseteq \mathcal{A}_g$ be an irreducible subvariety, not contained in any proper special subvariety of $\mathcal{A}_g$, such that, for every pair of positive integers i and n, we have that $Y \cap X_i^{(n)}$ is not Zariski dense in Y.
Then, there is a positive integer d_0 such that, for any $d \geq d_0$, "most" points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $X_i(\overline{\mathbb{Q}})$ for any positive integer i.

In this case our proof is closer in spirit to the Masser-Zannier one.
We are able to deduce the full analogue, i.e. allowing for Y being contained in a proper special subvariety of $\mathcal{A}_g$, conditionally on our form of isogeny estimates involving weakly special subvarieties, Proposition 3.8.14, to hold in $\mathcal{A}_g$.

Another different and substantially more complicated way of generalising the result of Masser-Zannier consists in finding "avoiding curves" rather than "avoiding points"; we state our main result in this sense, Theorem 4.2.1, where we require our variety to be sufficiently "asymmetric".

Theorem. *Let $C \subseteq Y(1)^3$ be a curve defined over $\overline{\mathbb{Q}}$ satisfying*

$$0 < 10^4 \deg_{x_1}(C) < 10^2 \deg_{x_2}(C) < \deg_{x_3}(C).$$

For "most" rational lines ℓ defined as

$$\{(t, at + b, ct + d) \in Y(1)^3 \mid t \in Y(1)\}$$

there is no point of $C(\overline{\mathbb{Q}})$ which is isogenous to any point of $\ell(\overline{\mathbb{Q}})$.

Above, we denote with x_1, x_2, x_3 the coordinates of $Y(1)^3$ respecting the product structure; the suitable notion of "most" for sets of rational lines is obtained by defining the height $h(\ell)$ of ℓ as

$$h(\ell) = \max\{h(a), h(b), h(c), h(d)\}.$$

1.2 Double unlikely intersections

Besides "avoiding problems", the second focus of this document is "double unlikely intersections"; this topic is relatively new, since the usual Zilber-Pink Conjecture does not take into account tangencies.
We quote the result by Marché-Maurin [36] dealing with curves in $\mathbb{G}_{\mathrm{m}}^2$ and the result by Corvaja-Demeio-Masser-Zannier [14] for a relative elliptic scenario.

The first of our results is a generalisation of the Theorem by Marché-Maurin. We quote our Theorem 5.0.1, obtained in collaboration with Capuano and Ottolini.

Theorem (B.-Capuano-Ottolini)**.** *Let n be a positive integer and let $C \subseteq \mathbb{G}_{\mathrm{m}}^n$ be a curve defined over $\overline{\mathbb{Q}}$. We suppose that C is not contained in any proper special subvariety of $\mathbb{G}_{\mathrm{m}}^n$; then there are only finitely many points $P \in C$ such that there is a one-codimensional special subvariety H satisfying:*

1. *P is in $C \cap H$;*
2. *the tangent space $T_P(C)$ of C at P is included in the tangent space $T_P(H)$ of H at P (namely, the curve C is tangent to H at P).*

Remark 1.2.1. *The Marché-Maurin result is the exact same statement as above in the restricted case of $n = 2$.*

The main difference with the original Marché-Maurin argument lies in the possibility of C being contained in a weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^n$: in this case, their argument, exploiting a classical upper bound on the height by Bombieri-Masser-Zannier (in [10]), needs to be replaced by ad hoc height estimates.
Our techniques use Puiseux series expansions and they might be of broader interest; indeed, our second result, concerned with curves in powers of $Y(1)$, follows from the same method, combined with Linear Forms in Logarithms (which is a result of Baker, in [7]) and the Pila-Wilkie Theorem 2.2.5.
We quote our Theorem 6.0.2, which really is just a translation from $\mathbb{G}_{\mathrm{m}}^n$ to $Y(1)^n$ of the statement above.

Theorem. *Let n be a positive integer and let $C \subseteq Y(1)^n$ be a curve defined over $\overline{\mathbb{Q}}$. We suppose that C is not contained in any proper special subvariety of $Y(1)^n$; then there are only finitely many points $P \in C(\overline{\mathbb{Q}})$ such that C is tangent to a one-dimensional special subvariety of $Y(1)^n$ at P.*

In both cases, namely in the context of powers of $\mathbb{G}_{\mathrm{m}}$ and powers of $Y(1)$, we are able to deduce corresponding "avoiding results" from the Theorems above.
We point out that such "avoiding results" are not "automatic" consequences of the "double unlikely intersections" statements: some book-keeping of the quantities arising during the various proofs is indeed required.

1.3 Structure of the document

After this Introduction, we start by recalling some basic notions and strategies behind unlikely intersections and the Pila-Zannier method, in chapter 2; the reader with an expertise in these topics can safely skip such content.
Chapter 3 contains the thematical core of the document: we focus on "avoiding problems", starting from simple examples, which are mainly intended to illustrate our techniques, then discussing the relevant generalisations of the Masser-Zannier result and finally concluding with the proofs of such.
In chapter 4 we focus on "higher dimensional avoiding problems", where we find "avoiding curves" rather than "avoiding points".
In chapter 5 we discuss "double unlikely intersections" in the context of powers of $\mathbb{G}_{\mathrm{m}}$, while in chapter 6 we discuss the same topic in the context of powers of $Y(1)$.
The last chapter 7 is devoted to results which are linked in some way to "avoiding problems" but that do not fall into that realm.

Chapter 2

Preliminaries

In this chapter we recall some basic notions of Logic and Number Theory, with the aim of introducing the main techniques and notions in the realm of point counting and unlikely intersections.
The reader with an experience in Diophantine Geometry can safely skip the whole chapter.

2.1 Definable sets

Our aim for this section is introducing semialgebraic sets, which will play a crucial role in the statement of the Pila-Wilkie Theorem.
The content below is classical in the context of Model Theory; we refer to the book [12].

Let us fix a language L, which is just a set of symbols that can either be constants, functions or relations, and an L-structure M, which is a set (that we will still denote with M) together with interpretations of the symbols of L, i.e. symbols of constants become elements of M, symbols of functions become functions from some power of M to M and symbols of relations become subsets of some power of M.
For instance, if L is the language of groups $\{*, ^{-1}, e\}$ then an L-structure M is a set together with a binary function $* : M^2 \to M$, a unary function $^{-1} : M \to M$ and an element $e \in M$ (note that these associations need not to respect the axioms of group theory). Of course groups with their usual interpretations of symbols are L-structures.

Using the symbols from L, the usual logical operators (i.e. $\neg, \wedge, \vee, \forall, \exists$) and the equality symbol $=$ we can formulate either sentences (which can be true or false in M) or formulae, which need to be evaluated at n-tuples of elements of M.
For instance, if L is the language of rings $\{+, -, *, 0, 1\}$ and φ is the sentence

$$\varphi : \forall x \exists y (y^2 = x)$$

then φ is true if M is a field with two elements but φ is false if M is a field with three elements. Instead, for a formula $\varphi(x_1, x_2)$ of the form

$$\varphi(x_1, x_2) : \exists y_1, y_2 ((y_1^2 = x_1) \wedge (y_2^2 = x_2))$$

its truth value depends on whether $x_1, x_2 \in M$ are squares.

Definition. *Given a language L and an L-structure M, a subset $S \subseteq M^n$ is definable (with parameters) if and only if there exist a formula $\varphi(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ and elements $c_1, c_2, \ldots, c_m \in M$ such that*

$$(a_1, a_2, \ldots, a_n) \in S \quad \textit{if and only if} \quad \varphi(a_1, a_2, \ldots, a_n, c_1, c_2, \ldots, c_m).$$

Example 2.1.1. *If L is the language of rings $\{+, -, *, 0, 1\}$ and we consider $\mathbb{R}$ as an L-structure, then the set $\{x \in \mathbb{R} \mid x \geq 0\}$ is definable using the formula $\varphi(x) : \exists y \mid x = y^2$. This is not in principle obvious, since $\geq$ is not a symbol in the language of rings!*

The role of $y_1, y_2, \ldots, y_m$ in the definition above is to allow parameters in our formulae; for instance, the subset $\{\pi\} \subseteq \mathbb{C}$ defined by $\varphi(x) : x - \pi = 0$ is not definable (without parameters) in the language of rings whenever π is irrational, but it is definable (with parameters) via $\varphi(x, y) : x - y = 0$.
We only stick to the notion of definability which allows parameters and the word "definable" for us will always mean "definable (with parameters)".

We now wish to study sets which are definable in the language of rings for the structures $\mathbb{C}$ and $\mathbb{R}$; these sets turn out to be the constructible sets and the semialgebraic sets respectively.

Remark 2.1.2. *A technical note: there is no need to distinguish between the language of rings and the language of fields, since the graph of the (partial) function for the inverse map*

$$^{-1} : M \setminus \{0\} \to M \setminus \{0\}$$

is definable as $\{(x, y) \mid xy = 1\}$.

A nice property for the definable sets of $\mathbb{C}$ is the so-called *quantifier elimination*. We say that an L-formula is *quantifier free* if it is obtained using only the symbols from L, the logical operators $\neg, \wedge, \vee$ and the equality symbol $=$; in particular, the quantifiers $\forall, \exists$ are not allowed in a quantifier free formula.

Theorem 2.1.3. *Let L be the language of rings and let $\varphi(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ be an L-formula. There exists a quantifier free L-formula $\psi(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ such that*

$$\varphi(a_1, a_2, \ldots, a_n, c_1, c_2, \ldots, c_m) \leftrightarrow \psi(a_1, a_2, \ldots, a_n, c_1, c_2, \ldots, c_m)$$

$$\textit{for every } (a_1, a_2, \ldots, a_n) \in \mathbb{C}^n \textit{ and } (c_1, c_2, \ldots, c_m) \in \mathbb{C}^m$$

where $\mathbb{C}$ is interpreted as an L-structure in the usual way (i.e. as a ring).

The subsets of $\mathbb{C}^n$ defined by a quantifier free formula are called *constructible sets*. These are the smallest collection $\mathcal{F}$ of subsets of $\mathbb{C}^n$ with the following properties:

1. for every polynomial $P(x_1, x_2, \ldots, x_n) \in \mathbb{C}[x_1, x_2, \ldots, x_n]$, the zero locus

$$\{(a_1, a_2, \ldots, a_n) \in \mathbb{C}^n \mid P(a_1, a_2, \ldots, a_n) = 0\}$$

   belongs to $\mathcal{F}$;

2. if $S \in \mathcal{F}$, then $\mathbb{C}^n \setminus S \in \mathcal{F}$ (this corresponds to the negation $\neg$ of the formula defining S);

3. if $S_1, S_2 \in \mathcal{F}$, then $S_1 \cap S_2 \in \mathcal{F}$ (this corresponds to the conjunction $\wedge$ of the two formulae defining S_1 and S_2);

4. if $S_1, S_2 \in \mathcal{F}$, then $S_1 \cup S_2 \in \mathcal{F}$ (this corresponds to the disjunction $\vee$ of the two formulae defining S_1 and S_2).

Every constructible set is therefore obtained by taking some subvarieties of $\mathbb{C}^n$ and applying to them intersections, unions and complements finitely many times.

Corollary 2.1.4. *The definable subsets of $\mathbb{C}^n$ (in the language of rings) are precisely the constructible sets.*

Notice that this means that constructible sets are stable by projection (which corresponds to $\exists$), since the projection of a definable set is definable. As a consequence, one obtains the following purely algebro-geometric result.

Theorem 2.1.5. *Let V be an algebraic variety over $\mathbb{C}$ and let $\varphi : V \to \mathbb{C}^n$ be a morphism over $\mathbb{C}$. The image of φ is a constructible subset of $\mathbb{C}^n$.*

Proof. The graph of φ is an algebraic subvariety $\Gamma(\varphi)$ of $V \times \mathbb{C}^n$ and the image of φ is the projection of $\Gamma(\varphi)$ on the second factor of $V \times \mathbb{C}^n$. □

Let us now shift our attention to $\mathbb{R}$. Here the situation is a bit more subtle: as we observed in Example 2.1.1, the set

$$\{x \in \mathbb{R} \mid x \geq 0\} = \{x \in \mathbb{R} \mid \exists y \in \mathbb{R} \text{ such that } y^2 = x\}$$

is definable and such a set cannot be obtained by applying intersections, unions and complements to subvarieties of $\mathbb{R}$; indeed, one obtains exactly the finite subsets of $\mathbb{R}$ and their complements.
However, it turns out that this is essentially the only "missing information" and we can circumvent it enlarging our language. Let $L = \{+, -, *, 0, 1, <\}$ be the language of ordered rings and we consider $\mathbb{R}$ as an L-structure by interpreting each symbol in the usual way.
Quantifier elimination holds in this larger language.

Theorem 2.1.6 (Tarski-Seidenberg Theorem)**.** *Let L be the language of ordered rings and let $\varphi(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ be an L-formula. There exists a quantifier free L-formula $\psi(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ such that*

$$\varphi(a_1, a_2, \ldots, a_n, c_1, c_2, \ldots, c_m) \leftrightarrow \psi(a_1, a_2, \ldots, a_n, c_1, c_2, \ldots, c_m)$$

$$\textit{for every } (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n \textit{ and } (c_1, c_2, \ldots, c_m) \in \mathbb{R}^m$$

where $\mathbb{R}$ is interpreted as an L-structure in the usual way (i.e. as an ordered ring).

The subsets of $\mathbb{R}^n$ defined by a quantifier free formula are called *semialgebraic sets*. We have again a nice description of these; they form the smallest collection $\mathcal{F}$ of subsets of $\mathbb{R}^n$ with the following properties:

1. for every polynomial $P(x_1, x_2, \ldots, x_n) \in \mathbb{R}[x_1, x_2, \ldots, x_n]$, its zero locus

   $$\{(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n \mid P(a_1, a_2, \ldots, a_n) = 0\}$$

   belongs to $\mathcal{F}$;

2. for every polynomial $P(x_1, x_2, \ldots, x_n) \in \mathbb{R}[x_1, x_2, \ldots, x_n]$, the set

   $$\{(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n \mid P(a_1, a_2, \ldots, a_n) > 0\}$$

   belongs to $\mathcal{F}$;

3. if $S \in \mathcal{F}$, then $\mathbb{R}^n \setminus S \in \mathcal{F}$ (this corresponds to the negation $\neg$ of the formula defining S);

4. if $S_1, S_2 \in \mathcal{F}$, then $S_1 \cap S_2 \in \mathcal{F}$ (this corresponds to the conjunction $\wedge$ of the two formulae defining S_1 and S_2);

5. if $S_1, S_2 \in \mathcal{F}$, then $S_1 \cup S_2 \in \mathcal{F}$ (this corresponds to the disjunction $\vee$ of the two formulae defining S_1 and S_2).

Corollary 2.1.7. *The definable subsets of $\mathbb{R}^n$ (in the language of ordered rings) are precisely the semialgebraic sets.*

Remark 2.1.8. *We observe that a subset of $\mathbb{R}^n$ is definable in the language of ordered rings precisely when it is definable in the language of rings, since the relation $<$ is already definable in the language of rings; $x < y$ if and only if*

$$\exists z \textit{ such that } y = z^2 + x \textit{ and } z \neq 0.$$

Again, semialgebraic sets are stable by projection (corresponding to $\exists$).
The semialgebraic subsets of $\mathbb{R}$ are exactly the finite unions of (possibly unbounded) intervals. This is immediate since any set of the form

$$\{x \in \mathbb{R} \mid P(x) > 0\} \text{ for some } P(x) \in \mathbb{R}[x]$$

is of such a shape. As a consequence, we obtain the following.

Theorem 2.1.9. *The definable subsets of* $\mathbb{R}$ *(in the language of ordered rings) are precisely the finite unions of (possibly unbounded) intervals.*

This is the assertion that $\mathbb{R}$, as an ordered ring, is an *o-minimal structure*. These structures are the main content of the next section.

2.2 O-minimal structures

In this section we focus on o-minimal structures, which are the structures for which the fundamental Pila-Wilkie Theorem 2.2.5 holds.

We now restrict our attention on the languages L including the symbols $\{+, -, *, 0, 1, <\}$ and possibly other symbols of function. We always consider $\mathbb{R}$ as an L-structure interpreting the symbols above in the usual way (i.e. as an ordered ring).

Definition. *We say that* $\mathbb{R}$ *is an o-minimal* L*-structure if every definable subset of* $\mathbb{R}$ *is a finite union of (possibly unbounded) intervals.*

Remark 2.2.1. *Theorem 2.1.9 above states precisely that* $\mathbb{R}$ *is o-minimal as an ordered ring.*

Notice that such property depends on L and on the interpretation of its symbols. For example, if we had a symbol $\sin \in L$ interpreted as the sine function, the set

$$\mathbb{Z} = \{x \in \mathbb{R} \mid \sin(\pi x) = 0\}$$

would be definable. As $\mathbb{Z}$ is not a finite union of intervals, the sine function is inadmissible in the o-minimal realm.

O-minimal structures are very rigid and definable sets are well understood (see for instance the notes of Starchenko [62]); they are described by the Cell Decomposition Theorem, which also implies that any definable set is a finite disjoint union of sets homeomorphic to some $\mathbb{R}^d$.

The language which is most relevant to us includes specifically:

- exp, a symbol of function of arity 1, which is interpreted as the real exponential function $\exp : \mathbb{R} \to \mathbb{R}$;
- f_α, symbols of function for each real analytic function $\alpha : [0,1]^d \to \mathbb{R}$ (defined on an open subset of $\mathbb{R}^d$ containing $[0,1]^d$) and interpreted as such function, with the convention that α is zero outside of $[0,1]^d$.

It has been proved by Van den Dries and Miller in [19] that this structure is o-minimal. When we refer to such a structure on the set $\mathbb{R}$, we use the notation $\mathbb{R}_{an,exp}$ to insist on the real exponentiation and the restricted analytic functions being allowed in the formulae.

Remark 2.2.2. *Since we are dealing with Diophantine Geometry questions, it might come as a surprise our need to work with* $\mathbb{R}$ *rather than with* $\mathbb{C}$*. The main reason is that we need to use crucially the Pila-Wilkie Theorem 2.2.3, which holds in the o-minimal world, while complex exponentiation would bring us much outside of such a realm.*
Indeed, the exponential function $\exp : \mathbb{C} \to \mathbb{C} \setminus \{0\}$ *would allow us to define*

$$\mathbb{Z} = \{z \in \mathbb{C} \mid \exp(2\pi i z) = 1\}$$

and, again, the definability of $\mathbb{Z}$ *prevents* $\mathbb{C}$ *together with* $\exp$ *from being an o-minimal structure (and, more in general,* $\mathbb{Z}$ *is very far from being a "tame" ring, for instance because of Gödel's Incompleteness Theorems).*
However, we can define a restricted complex exponentiation in $\mathbb{R}_{an,exp}$*, since (the graph of)*

$$e : [0,1) \times \mathbb{R} \to \mathbb{R}^2$$

$$e(a,b) = (\exp(-2\pi b)\cos 2\pi a, \exp(-2\pi b)\sin 2\pi a)$$

is definable. If we identify $\mathbb{C}$ *with* $\mathbb{R}^2$ *via the real and imaginary parts, this is precisely the graph of the function*

$$e(a+ib) = \exp(2\pi i(a+ib)) \text{ for } (a,b) \in [0,1) \times \mathbb{R}$$

restricted to the domain

$$\{z \in \mathbb{C} \mid 0 \leq \operatorname{Re} z < 1\}$$

and hence (the graph of) such restricted exponentiation is definable in $\mathbb{R}_{an,exp}$*.*
Notice that this domain itself is definable using inequalities, i.e. over $\mathbb{R}$ *rather than over* $\mathbb{C}$*. This is the general strategy behind the Pila-Zannier method, introduced in* [56]*, ubiquitous in this document: instead of working with the full uniformisers (like the full complex exponentiation) we restrict them to a "fundamental domain" (like the domain described above) in order to preserve o-minimality and then be able to apply the Pila-Wilkie Theorem.*

We are now ready to formulate the Pila-Wilkie Theorem. We first need the notion of *exponential height* of an n-tuple of rational numbers.

Definition. *The exponential height of an* n*-tuple of rational numbers* $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ *is*

$$H(\alpha_1, \alpha_2, \ldots, \alpha_n) = \max\{|p_1|, |q_1|, |p_2|, |q_2|, \ldots, |p_n|, |q_n|\}$$

where p_i *and* q_i *are coprime integers such that* $\alpha_i = p_i/q_i$ *for each* $1 \leq i \leq n$*.*

We give here a first so-called "blocks" version (see the paper of Pila [55]):

Theorem 2.2.3 (Pila-Wilkie, Pila)**.** *Let* L *be a language as above for which* $\mathbb{R}$ *is an o-minimal structure and let* $S \subseteq \mathbb{R}^n$ *be a definable set. For every positive real number* ε *there is a positive constant* C *such that, for every* $T \geq 1$*, the set*

$$\{P \in \mathbb{Q}^n \mid P \in S \text{ and } H(P) < T\}$$

is contained in a union of at most $C \cdot T^{\varepsilon}$ *connected semialgebraic sets contained in* S*.*

Remark 2.2.4. *We observe that some restriction on the definable sets of* $\mathbb{R}$ *is indeed necessary; for instance, being o-minimal provides such a reasonable constraint.*
If, say, $\mathbb{Z} \subseteq \mathbb{R}$ *was definable, then the conclusion above could not hold, as the size of the set*

$$\{P \in \mathbb{Q} \mid P \in \mathbb{Z} \text{ and } H(P) < T\}$$

is roughly $2T$ *for* T *big.*

We will often need a more generalised form of this Theorem, in two directions:

1. the set S would usually be allowed to vary in a *definable family* $S_{\underline{a}}$ and we ask for a uniform constant C not depending on the parameter $\underline{a}$;

2. the rational points of S would be points of "small" degree over $\mathbb{Q}$.

This is the ultimate version for us and we will stick to this for the whole document. We refer to the paper by Pila [51].

Theorem 2.2.5 (Pila-Wilkie, Pila)**.** *Let* L *be a language as above for which* $\mathbb{R}$ *is an o-minimal structure. Let* $S \subseteq \mathbb{R}^n$ *and* $A \subseteq \mathbb{R}^m$ *be definable sets, together with a function*

$$f : S \to A$$

whose graph is definable in $\mathbb{R}^n \times \mathbb{R}^m$*. For any* $\underline{a} \in A$*, we denote with* $S_{\underline{a}}$ *the fiber* $f^{-1}(\underline{a})$*.*

Let d *be a positive integer. For every positive real number* ε *there is a positive constant* C *such that, for every* $T \geq 1$ *and every* $\underline{a} \in A$*, the set*

$$\{P \in \overline{\mathbb{Q}}^n \mid P \in S_{\underline{a}}, H(P) < T \text{ and } [\mathbb{Q}(P) : \mathbb{Q}] < d\}$$

is contained in a union of at most $C \cdot T^{\varepsilon}$ *connected semialgebraic sets contained in* S*.*

For the notion of the exponential height of an n-tuple of algebraic points, we refer to section 2.6.

The Pila-Wilkie Theorem has been proven extremely useful in Diophantine Geometry, especially as a tool for the Pila-Zannier method. For an account of these developments, see [49].

Remark 2.2.6. *The proof of the Pila-Wilkie Theorem is not effective and most of our results, based on it, will not be effective as well. It is a nice challenge to try to prove some results without appealing to ineffective arguments and we are able to prove some simple results in this direction.*

2.3 A problem of Lang

We now show what is probably the most basic (nontrivial) application of the Pila-Wilkie Theorem. The scheme of the proof follows the now standard Pila-Zannier method, that we explain below.

In the 1960s Lang formulated the following statement, that was then proved true by Ihara, Serre and Tate with three independent proofs (see the account in the book of Zannier [67], which also treats this kind of issues in deep detail):

Theorem 2.3.1 (Lang, Ihara-Serre-Tate)**.** *Let $f(x,y)$ be an irreducible polynomial in $\mathbb{C}[x,y]$. Suppose there are infinitely many pairs of roots of unity $(\alpha_0,\beta_0),(\alpha_1,\beta_1),\dots$ such that $f(\alpha_i,\beta_i)=0$ holds for any $i\in\mathbb{N}$; then $f(x,y)$ is (up to constants) either of the form $x^a-\omega y^b$ or $x^ay^b-\omega$ for some coprime natural numbers a,b and a root of unity ω.*

We wish to rephrase this statement in the modern language of unlikely intersections.

Let us first recall that the *multiplicative group* $\mathbb{G}_{\mathrm{m}}$ is the algebraic group given by

$$\{(u,v)\in\mathbb{A}^2 \mid uv=1\}$$

with neutral element $(1,1)$ and operations $(u_1,v_1)\cdot(u_2,v_2)\to(u_1u_2,v_1v_2)$ and $(u,v)^{-1}\to(v,u)$. $\mathbb{G}_{\mathrm{m}}$ and $\mathbb{A}^1\setminus\{0\}$ are isomorphic as quasi-projective varieties and we will always think of $\mathbb{G}_{\mathrm{m}}$ as a subset of $\mathbb{A}^1$ in such a way (i.e. restricting to one coordinate).

The algebraic subgroups (i.e. Zariski closed subgroups) of $\mathbb{G}_{\mathrm{m}}^n$ are exactly the zero loci of systems of equations of the shape

$$\begin{cases} x_1^{p_{1,1}}\cdot x_2^{p_{2,1}}\cdot\ldots\cdot x_n^{p_{n,1}}=1\\ x_1^{p_{1,2}}\cdot x_2^{p_{2,2}}\cdot\ldots\cdot x_n^{p_{n,2}}=1\\ \dots\\ x_1^{p_{1,m}}\cdot x_2^{p_{2,m}}\cdot\ldots\cdot x_n^{p_{n,m}}=1 \end{cases}$$

where the $p_{i,j}$, for $1\le i\le n$ and $1\le j\le m$, are integers (this is a consequence of the basic theory of commutative complex Lie groups). Each connected component of an algebraic subgroup H is of the form $(\omega_1,\dots,\omega_n)\cdot H_0$ where the ω_i are suitable roots of unity and H_0 is the connected component of H passing through the origin.

Definition. *The special subvarieties of $\mathbb{G}_{\mathrm{m}}^n$ are the connected components of the algebraic subgroups; they are exactly the translates of irreducible algebraic subgroups by a torsion point.*
The special points of $\mathbb{G}_{\mathrm{m}}^n$ are the special subvarieties of dimension zero; they are exactly the torsion points, i.e. points of the form $(\omega_1,\dots,\omega_n)$ where each ω_i is a root of unity.

Notice that, with attention to the statement of Theorem 2.3.1 above:

1. $f(x,y)$ is, up to constants, of the form $x^a - \omega y^b$ or $x^a y^b - \omega$ (for some coprime natural numbers a, b and a root of unity ω) if and only if its zero locus is a connected component of an algebraic subgroup of $\mathbb{G}_{\mathrm{m}}^2$, i.e. a special subvariety;

2. $f(x,y)$ vanishes at infinitely many pairs of roots of unity (α_i, β_i), for $i \in \mathbb{N}$, if and only if its zero locus is a curve that contains infinitely many torsion points, i.e. special points.

A new reformulation of Theorem 2.3.1 using the "unlikely intersections language" just introduced is then:

Let X be a curve in $\mathbb{G}_{\mathrm{m}}^2$. If X contains infinitely many special points then X is a special curve.

Remark 2.3.2. *This is an André-Oort type statement.*

Simple proofs of this fact can be obtained by intersecting X with the zero loci of suitable auxiliary polynomials together with Bézout's Theorem or using heights (see [67] for an account of these proofs).

We want to show instead how to apply the Pila-Zannier method (and the Pila-Wilkie Theorem 2.2.3), with a proof which is very far in spirit from the ones existing before and which is much less elementary. The advantage is that it can be generalised to many other situations. We sketch how this method works:

1. we construct a suitable definable set S whose rational points correspond to the pairs of roots of unity lying on X;

2. we prove that S does not contain positive dimensional semialgebraic sets unless X is special;

3. we apply the Pila-Wilkie Theorem 2.2.3 to S exploiting 2. above and obtain an upper bound on the number of rational points on S;

4. we use Galois-theoretic considerations (i.e. the degree of the field of definition of roots of unity grows roughly like their order) and we obtain a lower bound on the number of rational points on S which contradicts the upper bound obtained in 3.

We now perform all these steps.

1. Let us fix the domain $D = \{z \in \mathbb{C} \mid 0 \leq \mathrm{Re}(z) < 1\}$ for the normalised exponential map $e : \mathbb{C} \to \mathbb{C} \setminus \{0\}$ given by $e(z) = \exp(2\pi i z)$; we observe that D is a fundamental domain for e, in the sense that e maps D to $\mathbb{C} \setminus \{0\}$ bijectively. Notice that the normalisation has been chosen so that

$$e(z) \text{ is a root of unity if and only if } z \text{ is a rational number.}$$

We consider the set $S \subseteq D^2$ (the preimage of X via e) given by

$$S = e^{-1}(X) \cap D^2 = \{(z_1, z_2) \in D^2 \mid f(e(z_1), e(z_2)) = 0\}.$$

If we identify $\mathbb{C}$ with $\mathbb{R}^2$ via the real and imaginary parts, we see that S is definable in $\mathbb{R}_{an,exp}$, since the normalised exponential e satisfies

$$e(a + ib) = \exp(-2\pi b)(\cos(2\pi a) + i \sin(2\pi a))$$

and of course (the graph of) $f(x, y)$ and D are semialgebraic sets; we just needed the (full) real exponentiation and the analytic functions cos and sin restricted to $[0, 2\pi]$.
We observe that e induces a bijection between the rational points of S and the pairs of roots of unity on X:

$$(z_1, z_2) \in S \leftrightarrow (e(z_1), e(z_2)) \in X \text{ for every } (z_1, z_2) \in D^2;$$

$$(z_1, z_2) \in \mathbb{Q}^2 \leftrightarrow (e(z_1), e(z_2)) \text{ is a pair of roots of unity, for every } (z_1, z_2) \in D^2.$$

2. We now would like to apply the Pila-Wilkie Theorem 2.2.3 to the set S, but we need first to classify the semialgebraic subsets of S.
Suppose that a positive dimensional semialgebraic variety is contained in S; in particular, a semialgebraic curve is contained in S, that is, there is an open interval $I \subseteq \mathbb{R}$ and real algebraic functions p_1, q_1, p_2, q_2 (that we can assume with no branches on I), not all constant, such that

$$(p_1(t) + iq_1(t), p_2(t) + iq_2(t)) \in S \text{ for every } t \in I$$

which is equivalent to

$$f(e(p_1(t) + iq_1(t)), e(p_2(t) + iq_2(t))) = 0 \text{ for every } t \in I.$$

We now observe that, forgetting about the condition $t \in I$, this last equations defines, at least locally, a complex analytic set (i.e. the zero locus of a holomorphic function of t). We can therefore find a suitable open enlargement $I \subseteq U \subseteq \mathbb{C}$ such that

$$f(e(p_1(t) + iq_1(t)), e(p_2(t) + iq_2(t))) = 0 \text{ for every } t \in U.$$

Remark 2.3.3. *A slightly annoying detail is that I might not always be enlarged to U since D is not open and we could actually "move outside of D", for instance when $p_1(t) = 0$ and $q_1(t) = t$; recall that the domain of e is indeed D. This issue can easily be solved choosing an open enlargement of D in advance; we do not pause on this.*

We try to give some intuition of what comes next. The equations above mean that we have found inside of S a non-empty (euclidean) open subset of a complex algebraic curve that, by definition of S, is mapped via e (essentially the exponential map) inside the algebraic curve given by the zero locus of $f(x, y)$.
This contradicts Ax's Theorem ([4]) unless something "special" is happening:

Theorem 2.3.4 (Ax)**.** *Let $u_1(t), u_2(t) \ldots, u_n(t)$ be holomorphic functions defined on an open subset $U \subseteq \mathbb{C}$, such that they are all independent over $\mathbb{Q}$ modulo constants (i.e., for no $q_1, q_2, \ldots, q_n \in \mathbb{Q}$, not all zero, the function $q_1u_1(t) + q_2u_2(t) + \ldots + q_nu_n(t)$ is constant). Then the field*

$$\mathbb{C}(t, u_1(t), u_2(t), \ldots, u_n(t), \exp(u_1(t)), \exp(u_2(t)), \ldots, \exp(u_n(t)))$$

has transcendence degree at least n over $\mathbb{C}(t)$.

Remark 2.3.5. *This Theorem roughly tells us that the only algebraic relations satisfied by the exponential function (in a functional context) are those coming from the identity*

$$\exp(a + b) = \exp(a) \cdot \exp(b).$$

The same is conjectured when the functions $u_1(t), u_2(t), \ldots, u_n(t)$ are replaced by complex numbers; this is the Schanuel's Conjecture, which is very wide open and far from being proven.

In our case, we set $n = 2$ and

$$u_1(t) = 2\pi i(p_1(t) + iq_1(t))$$

$$u_2(t) = 2\pi i(p_2(t) + iq_2(t))$$

which are both algebraic over $\mathbb{C}(t)$.
Since the image of S via e is contained in the algebraic curve defined by $f(x, y)$, then $\exp(u_1(t))$ and $\exp(u_2(t))$ are algebraically dependent and the transcendence degree of the relevant field is at most 1; by Ax's Theorem 2.3.4 this implies that $u_1(t)$ and $u_2(t)$ are $\mathbb{Q}$-linearly dependent modulo $\mathbb{C}$.
Seeing this linear dependence via e, we obtain that X must intersect some zero locus of equations either of the form $x^a y^b - \gamma$ or $x^a - \gamma y^b$ for natural numbers a, b (not both zero) and a nonzero complex number γ in an (infinite) open subset and thus X (which is irreducible) is itself that very zero locus. We observe that γ is then a root of unity (just plug one pair (α_n, β_n) into the equation).
These are precisely our exceptions, i.e. X is special as long as S contains a positive dimensional semialgebraic variety.

3. If X is not special, S does not contain any positive dimensional semialgebraic variety by the argument in 2. above. The Pila-Wilkie Theorem 2.2.3 then, for instance with $\varepsilon = 1/2$, implies:

   There is a positive real C such that, for every real $T \geq 1$, the rational points of S of exponential height at most T are less than $C \cdot T^{1/2}$.

4. We finally obtain a contradiction with the lower bound above by finding an upper bound given by Galois-theoretic considerations.
   Let K be the field of definition of X. First, $K \subseteq \overline{\mathbb{Q}}$: if not, let $X' \neq X$ be a Galois conjugate of X with respect to an automorphism of $\mathbb{C}$ which fixes $\overline{\mathbb{Q}}$; then $X \cap X'$ is infinite, as it contains all the pairs of roots of unity (α_i, β_i) for $i \in \mathbb{N}$, contradiction.
   Then $K \subseteq \overline{\mathbb{Q}}$ and therefore it is a number field (since K is finitely generated over $\mathbb{Q}$ using the finitely many coefficients of $f(x, y)$).
   So K is a finite extension of $\mathbb{Q}$ and let $d = [K : \mathbb{Q}]$. Let (α, β) be a pair of roots of unity on X. Let n be the maximum of the orders of α and β. The point (α, β) has at least n/d conjugates over K; indeed, at least a proportion of $1/d$ elements of $\mathrm{Gal}(\mathbb{Q}(\alpha, \beta), \mathbb{Q})$ fix a primitive element of K.
   All these conjugates lie on X, since X is defined over K. Therefore, we obtain at least $\varphi(n)/d$ points (α_i, β_i) (on X) for which α_i and β_i are roots of unity of order at most n. The distinct points $(e^{-1}(\alpha_i), e^{-1}(\beta_i))$ lie on S and they are rational numbers with exponential height at most n; the numerators are less than the denominators since they lie in $[0, 1)$. But then:

   the rational points of S of exponential height at most n are at least $\varphi(n)/d$.

   By the application of the Pila-Wilkie Theorem in 3. above, this implies

   $$\varphi(n)/d \leq C \cdot n^{1/2}$$

   which implies that n is bounded (since $\varphi(n) > n^c$ for any fixed $c < 1$ and big n). This implies finiteness of the special points and it completes the proofs.

2.4 Basics on elliptic curves

In this section we briefly recall some facts about elliptic curves; our main goal is the introduction of their (coarse) moduli space, the j-line $Y(1)$, which is the subject of the next section.

Each elliptic curve defined over $\mathbb{C}$ can be written as the projective completion of the affine plane curve

$$\{(x, y) \in \mathbb{A}^2 \mid y^2 = x^3 + ax + b\}$$

for some $a, b \in \mathbb{C}$ with $4a^3 + 27b^2 \neq 0$. The origin is intended to be the only point at infinity, namely the point $(0, 1, 0)$ of the projective completion

$$\{(X, Y, Z) \in \mathbb{P}_2 \mid Y^2Z = X^3 + aXZ^2 + bZ^3\}.$$

Viceversa, whenever $4a^3 + 27b^2 \neq 0$, the equations above define an elliptic curve.

The complex points of any elliptic curve E form a complex analytic manifold which is isomorphic (as a complex analytic manifold) to some quotient of $\mathbb{C}$

$$E(\mathbb{C}) \cong \mathbb{C}/\Lambda$$

for some *lattice* Λ, i.e.

$$\Lambda = \{p\omega_1 + q\omega_2 \mid p, q \in \mathbb{Z}\}$$

for some $\mathbb{R}$-linearly independent $\omega_1, \omega_2 \in \mathbb{C}$. Lattices are precisely the discrete rank 2 subgroups of $\mathbb{C}$. For any lattice $\Lambda \subseteq \mathbb{C}$, the quotient $\mathbb{C}/\Lambda$ corresponds to the complex points of some elliptic curve E over $\mathbb{C}$.
Remarkably, elliptic curves are endowed with an algebraic group law, meaning that there is a group morphism

$$+ : E \times E \to E$$

defined over $\mathbb{Q}(a, b)$ whenever E is the (projectivised) zero locus of the equation $y^2 = x^3+ax+b$. The correspondence between elliptic curves and quotients $\mathbb{C}/\Lambda$ for a lattice Λ gives rise to a group law on $\mathbb{C}/\Lambda$, which turns out to be simply the usual addition on $\mathbb{C}$ modulo Λ.

Morphisms between elliptic curves are very rigid.

Proposition 2.4.1. *Let E_1, E_2 be two elliptic curves over $\mathbb{C}$ and let*

$$\varphi : E_1 \to E_2$$

be a morphism over $\mathbb{C}$ which sends the origin of E_1 to the origin of E_2. Then φ is a group morphism.

From the point of view of lattices, for E_1 and E_2 corresponding to $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ respectively, any group morphism $\varphi : E_1 \to E_2$ gives rise to a $\mathbb{C}$-linear map

$$\varphi : \mathbb{C} \to \mathbb{C}$$

$$\varphi(z) = \gamma z \text{ for every } z \in \mathbb{C}$$

for some $\gamma \in \mathbb{C}$. Moreover, as φ is a function on $E_1(\mathbb{C}) \cong \mathbb{C}/\Lambda_1$ with values in $E_2(\mathbb{C}) \cong \mathbb{C}/\Lambda_2$, we have

$$\gamma\Lambda_1 \subseteq \Lambda_2.$$

Definition. *Any nonconstant group morphism between two elliptic curves E_1 and E_2 is called an isogeny.*

In the situation above, a group morphism is an isogeny precisely when $\gamma \neq 0$. The *degree* of an isogeny φ is equivalently:

1. the cardinality of the fiber $\varphi^{-1}(P)$ for any $P \in E_2$;
2. the index $[\Lambda_2 : \gamma\Lambda_1]$.

Remark 2.4.2. *Notice that for any lattice Λ and any positive integer n we have the inclusion $n\Lambda \subseteq \Lambda$, with $[\Lambda : n\Lambda] = n^2$. This situation corresponds to the multiplication-by-n map*

$$[n] : E \to E$$

on the elliptic curve E corresponding to Λ. This isogeny has degree n^2.

As an immediate consequence of the equivalence between 1. and 2. above, two elliptic curves over $\mathbb{C}$ are isomorphic if and only if they are associated to "homothetic" lattices (i.e. two lattices $\Lambda_1 \subseteq \mathbb{C}$ and $\Lambda_2 \subseteq \mathbb{C}$ are homothetic if and only if there is some nonzero $\gamma \in \mathbb{C}$ such that $\Lambda_2 = \gamma\Lambda_1$).
We therefore have the following one-to-one correspondence:

$$\{\text{Elliptic curves over } \mathbb{C} \text{ up to isomorphism over } \mathbb{C}\} \leftrightarrow \{\text{Lattices up to homothety}\}.$$

Remark 2.4.3. *It is very "rare" for two elliptic curves to be isogenous. The elliptic curves isogenous to a given one account only for countably many isomorphism classes (see e.g. Theorem 2.5.5), while the set of isomorphism classes of elliptic curves over $\mathbb{C}$ has continuum cardinality. This might be a prototypical avoiding problem: find an elliptic curve (or many) not isogenous to a given one. This problem is an immediate consequence of isogeny estimates, which are based on transcendental methods; we are not aware of a method (except maybe the arguments with bad reduction, but these fail if we add some "integrality condition"). We refer to the discussion of Masser and Zannier in* [42].

2.5 The moduli space of elliptic curves

We now introduce the *j-invariant* of an elliptic curve, a quantity which parameterises elliptic curves up to isomorphism.

Definition. *Let E be an elliptic curve over $\mathbb{C}$ given by the (projectivised) zero locus of the equation $y^2 = x^3 + ax + b$ for some $a, b \in \mathbb{C}$ such that $4a^3 + 27b^2 \neq 0$. The quantity*

$$j(E) = 1728 \cdot \frac{4a^3}{4a^3 + 27b^2}$$

is the j-invariant of E.

Theorem 2.5.1. *Two elliptic curves are isomorphic over $\mathbb{C}$ if and only if they have the same j-invariant.*

Remark 2.5.2. *An important observation is that if we wish to obtain a prescribed j-invariant we can always take a, b in its field of definition. In this sense the j-invariant gives a rational parameterisation: E can be chosen in its isomorphism class so that its field of definition is $\mathbb{Q}(j(E))$.*

Definition. *The j-line $Y(1)$ is the (coarse) moduli space of elliptic curves, meaning that it consists of a point for each isomorphism class of elliptic curves.*

Such a point is given by the j-invariant and hence $Y(1)$ is isomorphic to the affine line $\mathbb{A}^1$ as a quasi-projective variety. We identify $Y(1)(\mathbb{C})$ with $\mathbb{C}$ as a complex manifold.

Since any lattice gives rise to a unique isomorphism class of elliptic curves, it is natural to define a *j-function* of the lattice. We are also allowed to rescale our lattice by an homothety and we can always suppose our lattice to be of the form

$$\Lambda = \{p\tau + q \mid p, q \in \mathbb{Z}\}$$

for some $\tau \in \mathbb{H}$, where $\mathbb{H}$ is the *upper half-plane*

$$\mathbb{H} = \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\} \subseteq \mathbb{C}.$$

We define the j-function

$$j : \mathbb{H} \to \mathbb{C}$$

$$j(\tau) = \text{“the } j\text{-invariant of an elliptic curve corresponding to } \mathbb{C}/\Lambda\text{”}$$

$$\text{where } \Lambda = \{p\tau + q \mid p, q \in \mathbb{Z}\}.$$

Given $\tau_1, \tau_2 \in \mathbb{H}$, we have that

$$j(\tau_1) = j(\tau_2) \text{ if and only if}$$

$$\Lambda_1 = \{p\tau_1 + q \mid p, q \in \mathbb{Z}\} \text{ and } \Lambda_2 = \{p\tau_2 + q \mid p, q \in \mathbb{Z}\} \text{ are homothetic}$$

and we obtain the following result.

Proposition 2.5.3. *Let $\tau_1, \tau_2 \in \mathbb{H}$. We have that $j(\tau_1) = j(\tau_2)$ if and only if there are integers a, b, c, d such that $ad - bc = 1$ and*

$$\tau_2 = \frac{a\tau_1 + b}{c\tau_1 + d}.$$

Proof. Asking for Λ_1 and Λ_2 to be homothetic is the same as asking that some $\mathbb{Z}$-basis of Λ_1 is homothetic to a (fixed) $\mathbb{Z}$-basis of Λ_2.
If we choose $(\tau_2, 1)$ as the $\mathbb{Z}$-basis for Λ_2, then any $\mathbb{Z}$-basis of Λ_1 is exactly of the form $(a\tau_1 + b, c\tau_1 + d)$ whenever a, b, c, d are integers such that $ad - bc = \pm 1$. The result follows observing that

$$\operatorname{Im}\left(\frac{a\tau_1 + b}{c\tau_1 + d}\right) = \operatorname{Im}\tau_1 \cdot \frac{ad - bc}{|c\tau_1 + d|^2}$$

and thus $ad - bc$ must be positive whenever the left-hand side is. □

We can also construct a *fundamental domain* D for j, in the sense that the j-function gives a bijection, when restricted to D, to $\mathbb{C}$. The *standard fundamental domain* for the j-function is

$$D = \{\tau \in \mathbb{H} \mid -1/2 \leq \operatorname{Re}\tau < 1/2, |\tau| > 1\} \cup \{\tau \in \mathbb{H} \mid -1/2 \leq \operatorname{Re}\tau \leq 0, |\tau| = 1\}.$$

In the light of Proposition 2.5.3 above, one can prove that D contains exactly one point for each orbit of the action

$$\mathrm{SL}_2(\mathbb{Z}) \times \mathbb{H} \to \mathbb{H};$$

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tau\right) \to \frac{a\tau + b}{c\tau + d}.$$

Isogenies behave well with respect to the j-function. Let us take $\tau_1, \tau_2 \in \mathbb{H}$ with corresponding lattices

$$\Lambda_1 = \{p\tau_1 + q \mid p, q \in \mathbb{Z}\} \text{ and } \Lambda_2 = \{p\tau_2 + q \mid p, q \in \mathbb{Z}\}$$

and associated elliptic curves $E_1 = \mathbb{C}/\Lambda_1, E_2 = \mathbb{C}/\Lambda_2$. Observe that

E_1 and E_2 are related by an isogeny of degree n if and only if

there is some $\gamma \in \mathbb{C}$ such that $\gamma\Lambda_1 \subseteq \Lambda_2$ with $\Lambda_2/(\gamma\Lambda_1)$ of order n.

We say that an isogeny is *cyclic* if its kernel is a cyclic (finite) group; this is the same as the quotient $\Lambda_2/(\gamma\Lambda_1)$ above being cyclic. We obtain the following:

Proposition 2.5.4. *E_1 and E_2 are related by a cyclic isogeny of degree n if and only if there are integers a, b, c, d with $\gcd(a, b, c, d) = 1$, $ad - bc = n$ and*

$$\tau_2 = \frac{a\tau_1 + b}{c\tau_1 + d}.$$

Proof. Similar to that of Proposition 2.5.3, but with more hands-on work with lattices. □

We now define the *modular polynomials* $\Phi_n(x_1, x_2)$, which are one of the central objects of this document.

Theorem 2.5.5. *Let n be a positive integer. There is a unique irreducible polynomial $\Phi_n(x_1, x_2) \in \mathbb{C}[x_1, x_2]$ which is monic in x_1 such that*

$$\Phi_n(j(\tau_1), j(\tau_2)) = 0$$

if and only if E_1 and E_2 are related by a cyclic isogeny of degree n.

Definition. *The polynomials $\Phi_n(x_1, x_2)$ are called modular polynomials.*

Remark 2.5.6.

1. *Remarkably, the modular polynomials $\Phi_n(x_1, x_2)$ have integer coefficients.*
2. *The modular polynomials $\Phi_n(x_1, x_2)$ are symmetric (for $n \geq 2$). This amounts to the observation that, if the quotient $\Lambda_2/(\gamma\Lambda_1)$ is cyclic of order n, then also the quotient $\Lambda_1/(\gamma^{-1}n\Lambda_2)$ is cyclic of order n.*
3. *One can also define $\Phi_n(x_1, x_2)$ as the unique irreducible polynomial in $\mathbb{C}[x_1, x_2]$ which is monic in x_1 such that*

$$\Phi_n(j(\tau), j(n\tau)) = 0 \text{ for every } \tau \in \mathbb{H}.$$

Remark 2.5.7. *We give some motivation for the choice of cyclic isogenies rather than any general isogeny. There is indeed some "minimal" polynomial $\Phi'_n(x_1, x_2) \in \mathbb{C}[x_1, x_2]$ which is monic in x_1 such that*

$$\Phi'_n(j(\tau_1), j(\tau_2)) = 0$$

if and only if E_1 and E_2 are related by an isogeny of degree n, but it just turns out that

$$\Phi'_n(x_1, x_2) = \prod_{\substack{d|n \\ n/d \text{ is a square}}} \Phi_d(x_1, x_2)$$

since two elliptic curves E_1 and E_2 are related by an isogeny of degree n if and only if they are related by a cyclic isogeny of degree d for some d dividing n such that n/d is a square (this can be observed on lattices).
It is just more convenient to work directly with the irreducible factors $\Phi_d(x_1, x_2)$. Notice the analogy with the (arithmetic) situation in $\mathbb{G}_{\mathrm{m}}$, where $\Phi'_n(x_1, x_2)$ would correspond to the polynomial $x^n - 1$ and $\Phi_d(x_1, x_2)$ to a cyclotomic factor.

We can finally discuss some unlikely intersections notion on $Y(1)$.

Definition. *The special subvarieties of $Y(1)^n$ are exactly the irreducible components of a zero locus of a system of equations of the shape*

$$\begin{cases} \Phi_{k_1}(x_{a_1}, x_{b_1}) = 0 \\ \Phi_{k_2}(x_{a_2}, x_{b_2}) = 0 \\ \dots \\ \Phi_{k_m}(x_{a_m}, x_{b_m}) = 0 \end{cases}$$

where, for every $1 \leq i \leq m$, a_i, b_i are integers such that $1 \leq a_i, b_i \leq n$ and k_i is a positive integer.
The special points are the zero-dimensional special subvarieties.

Remark 2.5.8. *Notice that when $a_i = b_i$ the condition $\Phi_{k_i}(x_{a_i}, x_{b_i}) = 0$ amounts to a finite disjoint union of conditions of the form*

$$x_{a_i} = s_{i,h}$$

where $s_{i,h}$ is a singular modulus; we describe such notion right now.

Definition. *If $\tau \in \mathbb{H}$ is (imaginary) quadratic (i.e. $[\mathbb{Q}(\tau) : \mathbb{Q}] = 2$), we say that $j(\tau)$ is a singular modulus.*

Remark 2.5.9. *Suppose that E_1 and E_2 are isogenous; then $j(E_1)$ is a singular modulus if and only if $j(E_2)$ is.*

τ is a singular modulus if and only if any of the corresponding elliptic curves E such that

$$E(\mathbb{C}) \cong \mathbb{C}/\Lambda \text{ where } \Lambda = \{p\tau + q \mid p, q \in \mathbb{Z}\}$$

has "extra endomorphisms". The endomorphism ring $\mathrm{End}(E)$ is isomorphic to the ring

$$\{\gamma \in \mathbb{C} \mid \gamma\Lambda \subseteq \Lambda\}$$

essentially by definition (one can extract $\mathbb{C}$ from E seeing it as its Lie algebra - we do not pause on this) and this ring is seen to be either isomorphic to $\mathbb{Z}$ or to an order in a quadratic (imaginary) number field. In the latter case, E is said to have *CM* (Complex Multiplication).

Special points admit an easier description in terms of singular moduli.

Proposition 2.5.10. *A point $P \in Y(1)^n$ is special if and only if*

$$P = (s_1, s_2, \ldots, s_n)$$

for some singular moduli $s_1, s_2, \ldots, s_n$.

Remark 2.5.11. *In the unlikely intersections world, singular moduli in $Y(1)$ correspond to torsion points (i.e. roots of unity) in $\mathbb{G}_{\mathrm{m}}$; they are both special points.*

We conclude the section with a rather technical Lemma, which is a quantitative version of Proposition 2.5.4 with the additional hypothesis of τ_1, τ_2 restricted to the standard fundamental domain D.
This result turns out to be fundamental in the applications of the Pila-Zannier method to unlikely intersections problems in powers of $Y(1)$. Indeed, the coefficients $(a, b, c, d$ below) of a matrix corresponding to a "connecting isogeny" are often the coordinates of the rational points on a suitable definable set whose cardinality could be bounded by the Pila-Wilkie Theorem, provided we have an upper bound on $\max(|a|, |b|, |c|, |d|)$. We point out to Lemma 3.1 of [42] and [47] for a reference.

Lemma 2.5.12. *Let $\tau_1, \tau_2 \in D$, where D is the standard fundamental domain described above, such that for some positive integer n*

$$\Phi_n(j(\tau_1), j(\tau_2)) = 0.$$

Then there exist integers a, b, c, d such that $\gcd(a, b, c, d) = 1$ and $ad - bc = n$ satisfying:

$$\max(|a|, |b|, |c|, |d|) \leq n;$$

$$\tau_2 = \frac{a\tau_1 + b}{c\tau_1 + d}.$$

2.6 Basics on heights

We now define one of the main tools of Diophantine Geometry, the *height* of an algebraic number. Heights measure the "complexity" of an algebraic number; as a quantity, they are in a sense comparable to the number of bits required to store a given number in a computer. We refer to the survey on heights by Zannier [65].

Let us start with an elementary form of height. Given an algebraic number $\alpha \in \overline{\mathbb{Q}}$, there is a unique polynomial $P(t) \in \mathbb{Z}[t]$ such that:

1. $P(\alpha) = 0$;
2. $P(t)$ has minimal degree;
3. The leading coefficient of $P(t)$ is positive;
4. The coefficients $a_0, \ldots, a_n$ of $P(t) = a_0 + a_1 t + \ldots + a_n t^n$ are relatively prime integers.

Definition (Polynomial Height)**.** *The exponential polynomial height of α is*

$$H_{poly}(\alpha) = \max\{|a_0|, |a_1|, \ldots, |a_n|\}.$$

Thus, $H_{poly}(\alpha)$ is just the maximum of the absolute values of the coefficients of a minimal polynomial of α (over $\mathbb{Z}$).
One immediately obtains the following:

Proposition 2.6.1 (Northcott Property for the Polynomial Height)**.** *Let $C_1, C_2 \in \mathbb{R}$. The set*

$$\{a \in \overline{\mathbb{Q}} \mid [\mathbb{Q}(\alpha) : \mathbb{Q}] < C_1 \text{ and } H_{poly}(\alpha) < C_2\}$$

is finite.

Proof. Let $P(t)$ be a minimal polynomial for α as described above, such that $P(t) = a_0 + a_1 t + \ldots + a_n t^n$. Then n is bounded by C_1 and the values $|a_i|$ (for $i = 0, 1, \ldots, n$) are bounded by C_2, hence we obtain finiteness of the set of admissible minimal polynomials; their roots are the set above, which is finite. □

This Proposition (often called the *Northcott Property* of the height) allows us to count algebraic numbers of a given degree up to a certain height. This point of view will be crucial in the next sections, where we often prove results that hold for "most" algebraic numbers. The notion of "most" is indeed present in the main result of [42], which is indeed the Theorem that most part of this document is based on.

For our purposes it is more convenient to use a more refined height function, named *Weil height*, which has better functorial properties than the polynomial height and satisfies a useful local-global principle (the Weil height is indeed a product of *local heights*).

Let K be a number field. We start by studying the *absolute values* on K.

Definition. *An absolute value on K is a function*

$$|\cdot|_v : K \to \mathbb{R}_{\geq 0}$$

such that, for every $x, y \in K$:

1. $|x|_v = 0$ *if and only if* $x = 0$*;*
2. $|x + y|_v \leq |x|_v + |y|_v$ *(triangle inequality);*
3. $|xy|_v = |x|_v |y|_v$.

As an example, we have three types of distinct absolute values on $\mathbb{Q}$:

1. the trivial absolute value, given by $|x|_v = 1$ for every $x \neq 0$;
2. the usual archimedean absolute value $|x|_v = |x|$;
3. a rather more sophisticated class of absolute values, the *p-adic absolute values*. Given a prime number p, we set

$$|x|_v = \exp(-v_p(x)) \text{ for every } x \neq 0$$

where $v_p(x)$ is the p-adic evaluation of x (i.e. the exponent of p in the unique factorisation of x). Notice that a stronger form of the triangle inequality holds here

$$|x + y|_v \leq \max(|x|_v, |y|_v).$$

We observe that if $|\cdot|_v$ is an absolute value on K, then

$$|\cdot|_{v'} = |\cdot|_v^c \text{ for some } 0 < c < 1$$

is still an absolute value on K. This prompts us to give the following Definition:

Definition. *Two absolute values $|\cdot|_v$ and $|\cdot|_{v'}$ are said to be equivalent if there is some $c > 0$ such that $|\cdot|_{v'} = |\cdot|_v^c$. An equivalence class of absolute values is called a place.*

Theorem 2.6.2 (Ostrowski)**.** *Any absolute value on $\mathbb{Q}$ is equivalent to one described above*

We now generalise the situation to number fields. The main difference is the lack of unique factorisation of elements, but of course we use unique factorisation of ideals. We also wish to outline a single representative for each place.
We consider the following list of absolute values on K:

1. the trivial absolute value $|x|_v = 1$ for every $x \in K \setminus \{0\}$;

2. let $\sigma : K \to \mathbb{C}$ be an embedding of K into $\mathbb{C}$. Then we obtain an absolute value

$$|x|_v = |\sigma(x)| \text{ for every } x \in K$$

where $|\cdot|$ is the usual archimedean absolute value. These are called *archimedean absolute values* or *infinite places*.
We notice that the choices for σ are precisely $[K : \mathbb{Q}]$, but two embeddings σ and σ' which are conjugate, in the sense that

$$\sigma'(x) = \overline{\sigma(x)} \text{ for every } x \in K$$

where $\overline{z}$ denotes the complex conjugate of z, give rise to the same absolute value. One can show that the archimedean absolute values are exactly $r + s$, where r denotes the number of embeddings of K into $\mathbb{R}$ and s denotes the number of complex embeddings of K (so that $[K : \mathbb{Q}] = r + 2s$);

3. let $\mathfrak{p}$ be a prime ideal in the ring of integers O_K of K. We obtain an absolute value

$$|x|_v = |O_K/\mathfrak{p}|^{-v_\mathfrak{p}(x)} \text{ for every } x \in K \setminus \{0\}$$

where $v_\mathfrak{p}(x)$ is the exponent of $\mathfrak{p}$ in the unique factorisation of (x) (as a fractional ideal). These are called *non-archimedean absolute values* or *finite places*. As above, a stronger form of the triangle inequality holds; for this reason, these absolute values are often called *ultra-metric*.

We denote with $M_{K,inf}$ the set consisting of the single representatives of the infinite places of K as outlined above; analogously, we denote with $M_{K,fin}$ the set consisting of the single representatives of the finite places of K that we described before. Let $M_K = M_{K,inf} \cup M_{K,fin}$.

Theorem 2.6.3. *Any non-trivial absolute value over K is equivalent to an element of M_K.*

The normalisations we chose for the absolute values seem a bit arbitrary now; the following result motivates a bit our choice of representatives.

Theorem 2.6.4 (Product Formula)**.** *Let $x \in K \setminus \{0\}$. Then*

$$\prod_{v \in M_K} |x|_v = 1.$$

Remark 2.6.5. *The Product Formula is analogous to the fact that a rational function on a complete curve over $\mathbb{C}$ has the same number of zeroes and poles (both counted with multiplicities); indeed, the analogue of the height for a function field contained in $\overline{\mathbb{C}(t)}$ is a degree in the usual algebro-geometric sense.*

We can finally define the Weil height.

Definition (Weil height)**.** *Let $\alpha \in \overline{\mathbb{Q}}$ and let K be a number field such that $\alpha \in K$. The exponential Weil height of α is*

$$H(\alpha) = \left(\prod_{v \in M_K} \max\left(1, |x|_v\right) \right)^{1/[K:\mathbb{Q}]}$$

and the Weil height of α is

$$h(\alpha) = \log H(\alpha).$$

Remark 2.6.6. *Remarkably, the definitions above are independent of K.*

Notice that the Weil height comes with a decomposition into an infinite part H_{inf} (and h_{inf}) and a finite part H_{fin} (and h_{fin}), corresponding to infinite and finite places.
For instance, let p/q be a rational number, with p, q coprime integers. Then

$$H_{inf}(p/q) = \max\left(1, |p/q|\right)$$

$$H_{fin}(p/q) = |q|$$

and hence

$$H(p/q) = \max\left(|p|, |q|\right)$$

which is also a very reasonable notion of "complexity" of a rational number.

The cumbersome definition of the Weil height is compensated by a number of nice properties:

Proposition 2.6.7. *Let $\alpha, \beta \in \overline{\mathbb{Q}}$. Then:*

1. $h(\alpha + \beta) \leq h(\alpha) + h(\beta) + \log 2$;
2. $h(\alpha\beta) \leq h(\alpha) + h(\beta)$;
3. $h(\alpha^n) = |n|\, h(\alpha)$;
4. $h(\sigma(\alpha)) = h(\alpha)$ *for any* $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Proof. One just needs to check the inequalities at each single place. □

We stress that the definition of the Weil height as a product of local factors often allows one to deal with one place at a time and the proof of Proposition 2.6.7 is an example of how useful this local-global principle can be.

The Northcott Property holds as well.

Proposition 2.6.8 (Northcott Property for the Weil Height)**.** *Let $C_1, C_2 \in \mathbb{R}$. The set*

$$\{a \in \overline{\mathbb{Q}} \mid [\mathbb{Q}(\alpha) : \mathbb{Q}] < C_1 \text{ and } h(\alpha) < C_2\}$$

is finite.

Proof. Let $P(t) \in \mathbb{Q}(t)$ be the unique monic minimal polynomial for α and let $P(t) = q_0 + q_1 t + \ldots + q_{n-1}t^{n-1} + t^n$ for rational numbers $q_0, \ldots, q_{n-1}$. Then n is bounded by C_1. By 4. of the previous Proposition 2.6.7 all the Galois conjugates $\alpha_1, \alpha_2, \ldots, \alpha_n$ of α have the same height, bounded above by C_2. As q_i is a (fixed) symmetric polynomial evaluated in $\alpha_1, \alpha_2, \ldots, \alpha_n$, one obtains an upper bound on $h(q_i)$ using repeatedly 1. and 2. of the previous Proposition 2.6.7. Thus, the possibilities for the numerators and the denominators of each of $q_0, q_1, \ldots, q_{n-1}$ are finitely many and we conclude. □

The Weil height is the canonical height associated to the multiplicative group $\mathbb{G}_{\mathrm{m}}$ and the next Proposition is an example of this.

Proposition 2.6.9 (Kronecker)**.** *Let $x \in \overline{\mathbb{Q}} \setminus \{0\}$. Then*

$$h(x) = 0 \text{ if and only if } x \text{ is a root of unity.}$$

Proof. Suppose that $h(x) = 0$ and that x is not a root of unity. Then the sequence $x^1, x^2, x^3, \ldots$ is an infinite sequence of distinct algebraic numbers with height 0 (because of 3. of Proposition 2.6.7) and with degree bounded by $[\mathbb{Q}(x) : \mathbb{Q}]$; this contradicts the Northcott Property (Proposition 2.6.8).
Viceversa, if x is a root of unity of order n, again 3. of Proposition 2.6.7 implies

$$0 = h(1) = |n|\, h(x).$$

□

Corollary 2.6.10. *There are only finitely many roots of unity of a given degree over $\mathbb{Q}$.*

Proof. This is just a combination of the Northcott Property (Proposition 2.6.8) and the fact that roots of unity have zero height (Proposition 2.6.9). □

2.7 Heights and varieties

We are highly interested in the $\overline{\mathbb{Q}}$-points of algebraic varieties and we now define some height function on these. This is rather straightforward with all the machinery we developed already.

Let X be a quasi-projective variety defined over $\overline{\mathbb{Q}}$, together with an embedding (always defined over $\overline{\mathbb{Q}}$) $X \subseteq \mathbb{P}_n$ for some positive integer n (such an embedding exists precisely when X is quasi-projective). We put a height on $X(\overline{\mathbb{Q}})$ by simply putting a height on $\mathbb{P}_n(\overline{\mathbb{Q}})$. Thus, we just need the notion of Weil height for $\mathbb{P}_n(\overline{\mathbb{Q}})$.

Definition (Weil height on $\mathbb{P}_n$)**.** *Let $(x_0, x_1, \ldots, x_n) \in \mathbb{P}_n(\overline{\mathbb{Q}})$. Let K be a number field containing $\mathbb{Q}(x_0, x_1, \ldots, x_n)$. The exponential Weil height of $(x_0, x_1, \ldots, x_n)$ is defined as*

$$H(x_0, x_1, \ldots, x_n) = \prod_{v \in M_k} \max\left(|x_0|_v, |x_1|_v, \ldots, |x_n|_v\right).$$

The Weil height is defined as

$$h(x_0, x_1, \ldots, x_n) = \log H(x_0, x_1, \ldots, x_n).$$

Remark 2.7.1. *Again, the height is defined independently of K. Moreover, it is not in principle clear that the height is well-defined in $\mathbb{P}_n$, but*

$$H(x_0, x_1, \dots, x_n) = H(\lambda x_0, \lambda x_1, \dots, \lambda x_n) \text{ for every } \lambda \in \overline{\mathbb{Q}}$$

as a consequence of the product formula (Theorem 2.6.4). However, this prevents us from separating the infinite part and the infinite part of the height. This is an essentially unsolvable issue in the projective context: if we think of $\mathbb{P}_1(\overline{\mathbb{Q}})$ as the union of two copies A, B of $\overline{\mathbb{Q}}$, where we identify $A \setminus \{0\}$ and $B \setminus \{0\}$ via $t \to t^{-1}$, then a point $P \in P_1(\mathbb{Q})$ would correspond to some $p/q \in A$ and $q/p \in B$, for coprime integers p and q. The roles of the finite part and of the infinite part are reversed when moving from A to B.

The main property of the height is its behaviour with respect to morphisms.

Theorem 2.7.2 (Functoriality of the height)**.** *Let $X \subseteq \mathbb{P}_m$ and $Y \subseteq \mathbb{P}_n$ be quasi-projective varieties over $\overline{\mathbb{Q}}$ and let $\varphi : X \to Y$ a morphism, always defined over $\overline{\mathbb{Q}}$. We consider the Weil heights on X and Y given by their embeddings in $\mathbb{P}_m$ and $\mathbb{P}_n$ respectively.*

1. *There are two positive constants C_1, C_2 such that*

$$h(\varphi(P)) < C_1 h(P) + C_2 \text{ for every } P \in X(\overline{\mathbb{Q}}).$$

2. *If φ is quasi-finite (i.e. its fibers are finite), there are two positive constants C_1, C_2 and a proper (possibly reducible) subvariety $Z \subseteq X$ such that*

$$h(\varphi(P)) > C_1 h(P) - C_2 \text{ for every } P \in (X \setminus Z)(\overline{\mathbb{Q}}).$$

Proof. The upper bound estimate of 1. essentially amounts to an elaborate application of the basic properties of the height, namely 1. and 2. of Proposition 2.6.7; the idea is that "the image $\varphi(P)$ cannot be much more complicated than P" as, after all, φ is given by a fixed collection of polynomials. The proof for 2. is much deeper and it requires Hilbert's Nullstellensatz. We refer to chapter B.3 of the classical book by Hindry and Silverman [32]. □

Remark 2.7.3. *If one chooses "coherently" the two projective embeddings the constant C_1 of 2. can be taken to be exactly the degree d. A formal treatment of this would require some good understanding of line bundles; since we will not need these tools in the next sections, we do not pause on this.*

The general construction outlined above allows us to put a Weil height on any quasi-projective variety. In some cases, however, we might prefer a more direct approach; we now show with an example that they are usually equivalent.

Let K be a number field, let $a, b \in K$ such that $4a^3 + 27b^2 \neq 0$ and let E be the elliptic curve

$$E : \{(x, y) \in \mathbb{A}^2 \mid y^2 = x^3 + ax + b\} \cup \{\infty\}$$

that we think as embedded in $\mathbb{P}_2$ via $(x,y) \to (x,y,1)$ and $\infty \to (0,1,0)$, with image

$$E : \{(X,Y,Z) \in \mathbb{P}_2 \mid Y^2Z = X^3 + aXZ^2 + bZ^3\}.$$

We would like to read the height given by $E \subseteq \mathbb{P}_2$ off some more simple height. For instance, we are tempted to ignore the y-coordinate and set

$$h_E(x,y) = h(x) \text{ for every } (x,y) \in E(\overline{\mathbb{Q}}) \setminus \{\infty\} \text{ and } h_E(\infty) = 0$$

where the height $h(\cdot)$ on the right-hand side is the usual Weil height of $\overline{\mathbb{Q}}$ (i.e. the one of Definition 2.6). This is indeed compatible with the height of $\mathbb{P}_2$, thanks to the following Proposition.

Proposition 2.7.4. *There is a constant C, depending on a and b only such that:*

$$|h_{\mathbb{P}_2}(x,y) - (3/2)h(x)| < C \text{ for any } (x,y) \in E(\overline{\mathbb{Q}}) \setminus \{\infty\}$$

where $h_{\mathbb{P}_2}(x,y)$ is intended to be the height induced by $E \subseteq \mathbb{P}_2$ and $h(x)$ is the usual Weil height.

Proof. We split the proof in two parts:

1. We use 3. of Theorem 2.7.2 in order to find a positive constant C' such that

$$|h_{\mathbb{P}_2}(x,y) - C'h(x)| < C \text{ for any } (x,y) \in E(\overline{\mathbb{Q}}) \setminus \{\infty\}.$$

2. We prove that $C' = 3/2$ by taking x a very big positive integer tending to infinity.

1. We consider the morphism

$$\varphi : E \to \mathbb{P}_1$$

$$\varphi(x,y) = x \text{ and } \varphi(\infty) = \infty$$

and, as φ is a quasi-finite morphism (of degree 2) between smooth projective varieties, we apply 3. of Theorem 2.7.2. We obtain positive constants C_1 and C_2 such that

$$|h_{\mathbb{P}_1}(\varphi(x,y)) - C_1 h_{\mathbb{P}_2}(x,y)| < C_2 \text{ for every } (x,y) \in E(\overline{\mathbb{Q}}) \setminus \{\infty\}$$

with $h_{\mathbb{P}_1}(\varphi(x,y)) = h_{\mathbb{P}_1}(x,1)$ the height on $\mathbb{P}_1$ and $h_{\mathbb{P}_2}(x,y) = h_{\mathbb{P}_2}(x,y,1)$ the height on $\mathbb{P}_2$.
We now notice that, for a point $(\alpha,1) \in \mathbb{P}_1(\overline{\mathbb{Q}})$, the height $h_{\mathbb{P}_1}(\alpha,1)$ (in the sense of Definition 2.7) is exactly $h(\alpha)$, i.e. the usual Weil height in the sense of Definition 2.6. Then, taking $C' = 1/C_1$, there is some positive C such that

$$|h_{\mathbb{P}_2}(x,y) - C'h(x)| < C \text{ for any } (x,y) \in E(\overline{\mathbb{Q}}) \setminus \{\infty\}.$$

2. We now prove that $C' = 3/2$. Let x be a positive integer (that we think of as very big). There are constants δ_1, δ_2, independent of x, such that, for any $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$

$$\left|\left|\sigma(x^3 + ax + b)\right| - |x|^3\right| < \delta_1 \cdot |x| + \delta_2$$

$$\left||\sigma(y)|^2 - |x|^3\right| < \delta_1 \cdot |x| + \delta_2$$

$$\left||\sigma(y)| - |x|^{3/2}\right| < \delta_1 \cdot |x| + \delta_2$$

which implies, taking the "average" over all the conjugates of y, that for some positive constant δ_3 independent of x we have

$$|h_{inf}(y) - (3/2)h_{inf}(x)| < \delta_3 \cdot |x|^{-1/2}$$

where $h_{inf}(\cdot)$ denotes the infinite part of the Weil height.

We let L be a number field containing a, b, x, y (e.g. one can take $L = K(y)$) and write

$$[L : \mathbb{Q}]h_{\mathbb{P}_2}(x, y) = \sum_{v \in M_L} \log \max\left(|x|_v, |y|_v, 1\right) =$$

$$= \sum_{v \in M_{L,inf}} \log |y|_v + \sum_{v \in M_{L,fin}} \log \max\left(|x|_v, |y|_v, 1\right)$$

where the latter equality holds whenever x is a sufficiently big positive integer. For finite places $v \in M_{L,fin}$ we have that

$$|x|_v = 1$$

since x is a positive integer; we can write

$$|y|_v^2 = \left|x^3 + ax + b\right|_v \leq \max\left(|a|_v, |b|_v\right)$$

by the ultrametricity of the finite absolute values. Then we obtain bounds

$$h_{fin}(x) = 0 \text{ and } h_{fin}(y) \leq h_{fin}(a) + h_{fin}(b).$$

Using our equation for $[L : \mathbb{Q}]h_{\mathbb{P}_2}(x, y)$ we then obtain

$$|[L : \mathbb{Q}]h_{\mathbb{P}_2}(x, y) - [L : \mathbb{Q}]h_{inf}(y)| \leq h_{fin}(x) + h_{fin}(y) \leq \max\left(|a|_v, |b|_v\right)$$

and we therefore conclude by replacing $h_{inf}(y)$ with $(3/2)h_{inf}(x)$ and using the estimate above.

□

Remark 2.7.5. *It should be clear from the proof that the possibility of dealing separately with each single absolute value is one the main advantages of working with the Weil height, despite its complicated definition.*

One can also construct a "canonical" height $\hat{h}(\cdot)$ (often called the Néron-Tate height) such that the following analogue of Proposition 2.6.9 holds: for any $P \in E(\overline{\mathbb{Q}})$

$$\hat{h}(P) = 0 \text{ if and only } P \text{ is a torsion point.}$$

In this sense, the usual Weil height is already canonical for the multiplicative group $\mathbb{G}_{\mathrm{m}}$.

We now show some not immediately obvious application (Corollary 2.7.7) of the height machinery we just developed. Let E be an elliptic curve defined as the (projectivised) zero locus of the equation $y^2 = x^3 + ax + b$ for some $a, b \in K$ (such that $4a^3 + 27b^2 \neq 0$) for some number field K.
We consider the height $h_E(x, y) = h(x)$ discussed above, whenever $(x, y) \in E(\mathbb{Q}) \setminus \{\infty\}$.

Proposition 2.7.6. *There exists a positive constant C such that, for any torsion point $P \in E(\overline{\mathbb{Q}}) \setminus \{\infty\}$, we have $h_E(P) < C$.*

Proof. Let

$$[2] : E \to E$$

$$[2](P) = P + P \text{ for every } P \in E(\mathbb{C})$$

be the duplication map on E. Notice that [2] has degree 4 (e.g. since, for any lattice $\Lambda \subseteq \mathbb{C}$, we have $[\Lambda : (2\Lambda)] = 4$). We apply 2. of Theorem 2.7.2 and we obtain a positive constant δ_1 such that

$$|h_{\mathbb{P}_2}([2]P) - 4h_{\mathbb{P}_2}(P)| < \delta_1 \text{ for every } P \in E(\overline{\mathbb{Q}})$$

and, by Proposition 2.7.4, we can substitute $h_E(\cdot)$ to $h_{\mathbb{P}_2}(\cdot)$:

$$|h_E([2]P) - 4h_E(P)| < \delta_2 \text{ for every } P \in E(\overline{\mathbb{Q}})$$

for some positive constant δ_2. Suppose that P is a torsion point and that $h_E(P) > \delta_2/3$. Then we observe by induction that

$$h_E([2^k]P) \geq 4h_E([2^{k-1}]P) - \delta_2 > h_E([2^{k-1}]P) \geq \delta_2$$

for every positive integer k. The fact that P is torsion implies that there are two distinct positive integers $a < b$ such that

$$[2^a]P = [2^b]P$$

which contradicts

$$h_E[2^a]P < h_E[2^b]P.$$

Therefore, if P is a torsion point, we have

$$h_E(P) \leq \delta_2/3$$

and we conclude setting $C = \delta_2/3$. □

Corollary 2.7.7. *Let d be a positive integer. The set of torsion points $P \in E(\overline{\mathbb{Q}})$ such that*

$$[\mathbb{Q}(P) : \mathbb{Q}] < d$$

is finite.

Proof. This is an immediate application of the Northcott Property (Proposition 2.6.8) to the conclusion of the previous Proposition. □

Chapter 3

Avoiding problems

In this chapter we focus on some "avoiding problems" related to the Zilber-Pink Conjecture. The goal of these sections is proving a number of generalisations of Theorem 3.3.2 by Masser-Zannier. We point to our Theorem 3.6.7, as the main result of the chapter, for a statement in the setting of $Y(1)^r$.

3.1 The Zilber-Pink Conjecture

In this section we focus on one of the main conjectures in the realm of unlikely intersections, the Zilber-Pink Conjecture; for a treatment of this problem and the whole subject we refer to the book of Zannier [67] and the survey of Pila [49].

Let us compare the notions of "special subvariety" and "special point" that we have already introduced for powers of the j-line $Y(1)$ and for powers of the multiplicative group $\mathbb{G}_{\mathrm{m}}$ in the table below.

Ambient Variety.	$Y(1)^n$.	**Powers of** $\mathbb{G}_{\mathrm{m}}$.
Special subvarieties.	Irreducible components of the zero loci of systems of equations of the form $\Phi_k(x_i, x_j) = 0$.	Algebraic subgroups translated by a torsion point.
Special points.	Points of the form $(s_1, \ldots, s_n)$ where each s_i is a singular modulus.	Torsion points.

The main conjecture in the realm of unlikely intersections is the Zilber-Pink Conjecture, that we state here.

Conjecture (Zilber-Pink Conjecture)**.** *Let S be a Shimura variety and let $X \subseteq S$ be a subvariety which is not contained in any proper special subvariety of S. For a natural number k, let $S^{[k]}$ be the set of special subvarieties of S with codimension k. Then:*

$$\bigcup_{k \geq [1+\dim X],\ Y \in S^{[k]}} Y \cap X \text{ is not Zariski dense in } X$$

.

Remark 3.1.1. *The same is conjectured when S is a power of $\mathbb{G}_{\mathrm{m}}$ or an abelian variety.*

The philosophy behind this conjecture could be more or less that "Intersections of X with subvarieties with codimension greater than $\dim X$ are unlikely; if these subvarieties are allowed to vary in a countable family that comes from Diophantine Geometry, then either X itself is close to being special or X intersects the family sparsely".

One instance of this conjecture is the now proven (by Pila-Shankar-Tsimerman, with an appendix by Esnault-Groechenig, see [53]) André-Oort Conjecture:

Theorem 3.1.2 (André-Oort Conjecture)**.** *Let S be a Shimura variety. The closure of any set of special points of S is a finite union of special subvarieties of S.*

Remark 3.1.3. *The same result has been proven for S either a power of $\mathbb{G}_{\mathrm{m}}$ or an abelian variety.*

For instance, if S is a power of $\mathbb{G}_{\mathrm{m}}$ (or an abelian variety) and $X \subseteq S$ is a subvariety that contains no torsion translate of an algebraic subgroup of S, then X contains only finitely many torsion points. This special case is also known as Manin-Mumford.

The Zilber-Pink Conjecture is wide open in general. For example, we do not have full proofs of the following:

Conjecture. *Let $X \subseteq Y(1)^3$ be a curve. The set*

$$\{(x_1, x_2, x_3) \in X(\mathbb{C}) \mid \Phi_m(x_1, x_2) = 0 \text{ and } \Phi_n(x_2, x_3) = 0 \text{ for some } m, n\}$$

is finite, unless X is contained in the zero locus of some $\Phi_k(x_i, x_j)$ (i.e. in a special surface).

Partial cases have been proven by Habegger and Habegger-Pila for asymmetric curves in [29] and [30] and a recent major breakthrough has been made by Daw-Orr for curves with a prescribed behaviour at infinity [17]; this last paper is similar to our arguments for a curve tangent to special subvarieties in a power of $Y(1)$.

In the case of powers of $\mathbb{G}_{\mathrm{m}}$ and abelian varieties, the Zilber-Pink Conjecture has been proven for curves by Bombieri-Masser-Zannier and Maurin in the case of powers of $\mathbb{G}_{\mathrm{m}}$ and by Habegger-Pila in the case of abelian varieties.

Theorem 3.1.4. *The Zilber-Pink Conjecture holds for curves in powers of $\mathbb{G}_{\mathrm{m}}$ (or in an abelian variety).*

The case of higher dimensional varieties is still open.

One can see that the condition on dimensions in the Zilber-Pink Conjecture is the correct one and indeed Eterović and Scanlon [21] proved "Likely Intersections" results under suitable hypotheses, even replacing Zariski density with euclidean density.

Theorem 3.1.5 (Eterović-Scanlon, Likely Intersections)**.** *Let S be a Shimura variety and let $X \subseteq S$ be a subvariety "satisfying suitable conditions". Then the set*

$$\bigcup_{Y \in S^{[\dim X]}} Y \cap X$$

is euclidean dense in X.

3.2 Some related conjectures

In this section we discuss two conjectures which would imply the Zilber-Pink Conjecture 3.1 in a number of cases.

The first conjecture, which is wide open, that we mention is the "Strong Galois Orbit" Conjecture (formulated, for instance, by Pila in [50]):

Conjecture (SGO)**.** *Let $C \subseteq Y(1)^2$ be a non-special curve over $\overline{\mathbb{Q}}$ and let $Y_n = \{(x,y) \mid \Phi_n(x,y) = 0\}$ be the modular curves. Then there are positive constants c, δ such that*

$$\textit{If } (x,y) \in C \cap Y_n \textit{ then } [\mathbb{Q}(x,y) : \mathbb{Q}] \geq c \cdot n^{\delta}$$

.

Remark 3.2.1. *This conjecture is believed to be true even without mentioning C (this is discussed in the aforementioned paper by Pila).*

It follows from a work of Habegger-Pila [30] that SGO implies Zilber-Pink for curves in any power of $Y(1)$. Many of our results will be proven obtaining a form of SGO, of course with additional hypotheses.
A related (and stronger) conjecture is the weakly bounded height Conjecture; this has been formulated by Habegger in [29]. We formulate two versions of different strength.

Conjecture (Weakly bounded height)**.** *Let $C \subseteq Y(1)^2$ be a non-special curve over $\overline{\mathbb{Q}}$ and let $Y_n = \{(x,y) \mid \Phi_n(x,y) = 0\}$ be the modular curves. Then there are positive constants c, δ (with $\delta < 1/2$) such that:*

1. $$\textit{If } (x,y) \in C \cap Y_n \textit{ then } h(x,y) \leq c \cdot n^{\delta}$$

   ;

2. $$\textit{If } (x,y) \in C \cap Y_n \textit{ then } h(x,y) \leq c \cdot \log n$$

   .

We prove the stronger weakly bounded height Conjecture 2. with the hypothesis of C being tangent to Y_n at (x, y); actually we use this in order to prove finiteness of such possible n and (x, y), so a posteriori the result is vacuous.

The weakly bounded height Conjecture 1. implies SGO; this is immediate via the so-called isogeny estimates. These include a number of results obtained first by Masser-Wüstholz [41] and refined by Gaudron-Rémond [23]. We give the Gaudron-Rémond version.

Theorem 3.2.2 (Isogeny estimates)**.** *There is an effective positive constant c with the following property. Suppose that $x, y \in \overline{\mathbb{Q}}$ are defined over a number field of degree $d \geq 2$ over $\mathbb{Q}$ and that $\Phi_n(x, y) = 0$ for some positive integer n. Then there exists n' such that*

$$\Phi_{n'}(x, y) = 0 \text{ and } n' \leq c \cdot h(x)^2 \cdot d^2 \log^2 d$$

.

Remark 3.2.3. *Notice that n' does not depend on $h(y)$; this is somewhat a special feature of the modular setting. In the multiplicative/abelian context a rational point can have arbitrary "big" rational multiples and their height must then be taken into account for an estimate of n'. This result of Masser is the correct analogue* [38].

Remark 3.2.4. *If none (or at most one) among x, y is a singular modulus, then $n' = n$.*

Isogeny estimates are obtained with transcendental methods, which give very uniform bounds, but usually without an optimal dependency on the degree (in the sense that we show below). If one uses Serre's Open Image theorem, for instance, one gets a very strong form of SGO for a fixed x.

Let $x \in \overline{\mathbb{Q}}$ be fixed and not a singular modulus. There is a positive constant L such that:

$$\Phi_n(x, y) = 0 \text{ implies } [\mathbb{Q}(x, y) : \mathbb{Q}] \geq n/L.$$

L has been now made effective but its dependence on x is bad (at least for our purposes). See the paper of Lombardo [35]. Notice that with isogeny estimates one would only obtain a lower bound of the form

$$\Phi_n(x, y) = 0 \text{ implies } [\mathbb{Q}(x, y) : \mathbb{Q}] \geq c' \cdot n^{1/2} / \log n$$

for some effective universal positive constant c'.

3.3 Avoiding problems and literature

We now discuss a Theorem of Masser and Zannier, whose generalisations and variations will be the main content of this document.
Recently, Masser and Zannier proved a result which is close in spirit to unlikely intersections,

but of a rather different nature than the Zilber-Pink Conjecture 3.1. They proved that "most" abelian varieties are not isogenous to the jacobian of a curve. See [42].

Let us describe the meaning of "most" here. We consider some height function (e.g. the Weil height of Definition 2.6) $h : \overline{\mathbb{Q}} \to \mathbb{R}$ with the Northcott Property (Proposition 2.6.8):

For any positive integer d and any real number T the set

$$\{\alpha \in \overline{\mathbb{Q}} \mid [\mathbb{Q}(\alpha) : \mathbb{Q}] \leq d, h(\alpha) < T\}$$

is finite.

Let $\mathcal{S}$ be a set of algebraic numbers and let d be a fixed positive integer. We say that a property (P) holds for "most" numbers of $\mathcal{S}$ (of degree $\leq d$) if:

$$\lim_{T \to +\infty} \frac{|\{\alpha \in \mathcal{S} \mid h(\alpha) < T, [\mathbb{Q}(\alpha) : \mathbb{Q}] \leq d \text{ and } \alpha \text{ satisfies (P)}\}|}{|\{\alpha \in \mathcal{S} \mid h(\alpha) < T, [\mathbb{Q}(\alpha) : \mathbb{Q}] \leq d\}|} = 1.$$

The same notion of "most" is also intended for $\overline{\mathbb{Q}}$-points of an algebraic quasi-projective variety.

In section 2.5 we have seen that elliptic curves are parameterised by the (coarse) moduli space $Y(1)$, in the sense that the points of $Y(1)(\mathbb{C})$ correspond to the isomorphism classes of elliptic curves over $\mathbb{C}$ and, moreover, for any field $K \subseteq \mathbb{C}$ and isomorphism class $p \in Y(1)(\mathbb{C})$

$$p \text{ admits a model over } K \text{ if and only if } p \in Y(1)(K).$$

Similar properties hold for higher dimensional abelian varieties.

Definition. *Let $g \geq 1$ be a positive integer. The (coarse) moduli space of abelian varieties of dimension g is $\mathcal{A}_g$.*

As above, points of $\mathcal{A}_g(\mathbb{C})$ correspond to the isomorphism classes of abelian varieties over $\mathbb{C}$ and for any field $K \subseteq \mathbb{C}$ and isomorphism class $p \in \mathcal{A}_g(\mathbb{C})$

$$p \text{ admits a model over } K \text{ if and only if } p \in \mathcal{A}_g(K).$$

It turns out that $\mathcal{A}_g$ is (if we forget about its moduli space nature) a quasi-projective variety of dimension $\frac{g(g+1)}{2}$.

Definition. *We say that two abelian varieties A_1, A_2 over $\mathbb{C}$ are isogenous if there exists a morphism (over $\mathbb{C}$)*

$$\varphi : A_1 \to A_2$$

such that φ is surjective and the kernel of φ is finite.
We say that two points $p, q \in \mathcal{A}_g(\mathbb{C})$ are isogenous if the corresponding abelian varieties are isogenous.

Remark 3.3.1. *The morphism φ above is automatically a group morphism.*

We can finally state the first "avoiding problem", which is the main Theorem of [42].

Theorem 3.3.2 (Masser, Zannier)**.** *Let $g \geq 4$ and let $\mathcal{H} \subseteq \mathcal{A}_g$ be a proper subvariety. For any $d \geq C(g)$, where $C(g)$ is an explicit constant depending on g only, we have that "most" points of $\mathcal{A}_g(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $\mathcal{H}(\overline{\mathbb{Q}})$.*

Remark 3.3.3. *Let $\mathcal{H}$ be the Torelli locus, i.e. the subset of $\mathcal{A}_g$ corresponding to the jacobians of curves of genus g. When $g \geq 4$, the "generic" abelian variety is not a jacobian since $\mathcal{H}$ has dimension $3g - 3 < \frac{g(g+1)}{2} = \dim \mathcal{A}_g$.*
Theorem 3.3.2 above tells us that "most" points of $\mathcal{A}_g$ are not isogenous to any point in the Torelli locus, i.e. "most" abelian varieties are not isogenous to a jacobian, even over $\overline{\mathbb{Q}}$. This was the first instance where "general" abelian varieties (as opposed to, say, CM ones) were exhibited not to be isogenous to jacobians of curves.

Remark 3.3.4. *Notice that if we replace $\overline{\mathbb{Q}}$ with $\mathbb{C}$ then the statement becomes immediate. Given a proper subvariety $\mathcal{H} \subseteq \mathcal{A}_g$ and a positive integer n, the subset of $\mathcal{A}_g$ consisting of the points isogenous to some point in $\mathcal{H}$ via an isogeny of degree n is (contained in) a proper subvariety $\mathcal{H}_n$. Hence $\mathcal{A}_g$ cannot be covered by a countable union of proper subvarieties $\mathcal{H}_1, \mathcal{H}_2, \ldots$ by measure-theoretic reasons. This is by no means obvious over $\overline{\mathbb{Q}}$.*

A similar statement has been considered by Tsimerman [63], who does consider not only isogenies but also quotients; for related questions in positive characteristic see also the paper by Shankar-Tsimerman [61] and especially the paper by Asvin [3] who asks about "a curve avoiding a curve".

3.4 First basic results and proofs

In this section we outline the proofs of a number of results, with the main purpose of developing the main techniques we will use later.
Besides the isogeny estimates Theorem 3.2.2 and the Pila-Wilkie Theorem 2.2.5, that we have already seen, we will encounter some new functional transcendence results of the Ax-Schanuel type, concerning $Y(1)$ rather than $\mathbb{G}_{\mathrm{m}}$ as in Theorem 2.3.4: this is a result obtained by Pila-Tsimerman in [54], which is our Theorem 3.4.9.

It will be convenient for us to denote with Q_T the set

$$\{q \in Y(1)(\mathbb{Q}) \mid h(q) < \log T\}.$$

A simple version (for $Y(1)^2$ instead of $\mathcal{A}_g$) of the Masser-Zannier result is given by Pila in his book [52].

Proposition 3.4.1 (Masser-Zannier, Pila)**.** *Let $C \subseteq Y(1)^2$ be a curve over $\overline{\mathbb{Q}}$. There exists a constant c such that, for every real number $T \geq 2$, the number of points of Q_T^2 isogenous to some point of C is at most $c \cdot T^2 (\log T)^{37}$.*

Remark 3.4.2. *The constant c above is not effective, since it depends on the ineffective Pila-Wilkie Theorem 2.2.5.*

Remark 3.4.3. *The Proposition is immediate whenever C is not defined over $\overline{\mathbb{Q}}$.Indeed, let $\sigma \in \mathrm{Gal}(\mathbb{C}/\overline{\mathbb{Q}})$ which acts nontrivially on C. Then $C \cap \sigma(C)$ consists of finitely many points and these include all the algebraic points of C. One then concludes using Lemma 3.4.7 below.*

Remark 3.4.4. *Notice that the number of points of Q_T^2 is roughly $(144/\pi^4)T^4$.*

Proof. We need to count (twice) the number of pairs (a, b) of relatively prime positive integers such that $a, b \leq T$. Such quantity is

$$\sum_{\substack{a,b \leq T \\ \gcd(a,b)=1}} 1 = 2 \sum_{\substack{a \leq b \leq T \\ \gcd(a,b)=1}} 1 = 2 \sum_{b \leq T} \varphi(b) = 2 \sum_{b \leq T} \sum_{d|b} \left(\mu(d) \frac{b}{d} \right) = 2 \sum_{d \leq T} \sum_{c \leq \lfloor T/d \rfloor} (\mu(d)c) =$$

$$= 2 \sum_{d \leq T} \mu(d) \left(\frac{1}{2} \lfloor T/d \rfloor^2 + \frac{1}{2} \lfloor T/d \rfloor \right) = \sum_{d \leq T^{2/3}} \mu(d) \lfloor T/d \rfloor^2 + \sum_{T^{2/3} < d \leq T} \mu(d) \lfloor T/d \rfloor^2 +$$

$$+ \sum_{d \leq T} \mu(d) \lfloor T/d \rfloor = \sum_{d \leq T^{2/3}} \mu(d) \frac{T^2}{d^2} + C_1 \text{ with } |C_1| \leq 3T^{5/3}$$

and from

$$\sum_{d=1}^{N} \mu(d) \frac{1}{d^2} = \zeta(2)^{-1} + C_2 = \frac{6}{\pi^2} + C_2 \text{ with } |C_2| \leq \frac{1}{N}$$

the result follows. □

We also show how to improve the bounds in a concrete case, i.e. when C is given by a unit equation.

Proposition 3.4.5. *Let $C = \{(x, y) \in Y(1)^2 \mid y = x + 1\}$. There is a finite union Z of curves and a constant c such that, for every real number $T \geq 2$, the number of points of $Q_T^2 \setminus Z(\mathbb{Q})$ isogenous to some point of C is at most $c \cdot (\log T)^{200}$.*

Remark 3.4.6. *The number of points of $Z(\mathbb{Q}) \cap Q_T^2$ is bounded by $c' \cdot T^2$ for some constant c' independent of T. This essentially cannot be avoided: the intersection $C(\mathbb{Q}) \cap Q_T^2$ already has cardinality bounded below by T^2 for T big enough and of course any point of $C(\mathbb{Q})$ is isogenous to some point of C (e.g. the point itself).*

We prove both results at once, using a direct translation of the methods used by Masser-Zannier for their Theorem 3.3.2. We will use the Pila-Zannier method. We divide the proof in a few steps:

1. first, we construct a definable set $S_{p,q}$ with the property that integral points of $S_{p,q}$ correspond to isogenies between $(p, q) \in Q_T^2$ and a point $(x, y) \in C(\mathbb{C})$;

2. we then classify the semialgebraic varieties contained in $S_{p,q}$ using an analogue of Ax's Theorem 2.3.4. In this case, in contrast to the proof in section 2.3, we can actually have some positive dimensional connected semialgebraic varieties contained in $S_{p,q}$, but we show that they account for a single couple $(x, y) \in C(\mathbb{C})$;

3. we apply the Pila-Wilkie Theorem on $S_{p,q}$ and we obtain an upper bound on the number of rational points with bounded height (in terms of T);

4. we use isogeny estimates in order to prove that if (p, q) and (x, y) are isogenous then $[\mathbb{Q}(x, y) : \mathbb{Q}]$ is "big" in terms of the degrees of the connecting isogenies. We can therefore use Galois conjugates and produce many rational points (accounting for distinct $(x, y) \in C(\mathbb{C})$) on $S_{p,q}$ and in turn we deduce (using the conclusion of 3.) an upper bound on the degrees of the connecting isogenies in terms of T only;

5. we prove that the "isogenous images" of $C(\mathbb{C})$ via the isogenies (of bounded degree by 4.) are far from covering Q_T^2 and thus "most" points of Q_T^2 are not isogenous to any point of $C(\mathbb{C})$. This concludes the proof of Proposition 3.4.1;

6. we improve our bound in the case of a unit equation. The main idea here is that C will never be contained in its "isogenous images" and we bound the number of points in the intersections with these rather than counting in Q_T^2.

We have not explicitly defined yet what we mean for two points $(p, q), (x, y) \in Y(1)(\mathbb{C})$ being isogenous. We can recall the following notation:

$$\text{we say that } (p, q) \text{ and } (x, y) \text{ are isogenous if the abelian varieties } E_p \times E_q \text{ and } E_x \times E_y \text{ are isogenous}$$

whenever E_p, E_q, E_x, E_y are elliptic curves with j-invariants p, q, x, y respectively.

A quick helpful observation follows from the semisimplicity of abelian varieties up to isogenies: if (p, q) is isogenous to (x, y), then either

$$p \text{ is isogenous to } x \text{ and } q \text{ is isogenous to } y$$

or

$$p \text{ is isogenous to } y \text{ and } q \text{ is isogenous to } x.$$

We always stick to the first case, obtaining the full result simply repeating the argument for the "symmetric curve" C' defined as

$$C' = \{(x, y) \in Y(1)^2 \mid (y, x) \in C\}.$$

In this sense, we can assume isogenies always "respect factors".
Before we step into the proof, we get rid of a few "degenerate" cases, i.e. the cases of C weakly special. We point to section 3.8 for a general description of weakly special subvarieties of powers of $Y(1)$.

Lemma 3.4.7. *There is a positive constant c such that, for every real $T \geq 2$ and every $p \in Q_T$, the set consisting of the $q \in Q_T$ which are isogenous to p has size at most*

$$c \cdot (\log T)^5.$$

Proof. First notice that isogeny estimates (Theorem 3.2.2) imply that if $p, q \in Q_T$ are isogenous, then there is some n' such that

$$\Phi_{n'}(p, q) = 0$$

$$n' \leq c_0 \cdot h(p)^2 \cdot 2^2 \log^2 2 \leq c_1 \cdot (\log T)^2$$

for some effective universal constant c_0, i.e. p and q are related by an isogeny of degree at most $c_1 \cdot (\log T)^2$, where c_1 is a universal (explicit) constant.

Given a positive integer m, the degree of $\Phi_m(x, y)$ in each coordinate is precisely $\psi(m)$, where ψ is the Dedekind ψ function (i.e. $\psi(m)$ is the sum of the divisors d of m such that m/d is squarefree). This can be shown by counting the number of sublattices (of a given one) with cyclic quotient of order n.
An application of a classical Theorem of Mertens [45] shows that there exists a constant c_2 such that $\psi(m) \leq c_2 \cdot m \log m$ for every $m \geq 2$:

$$\psi(m)/m = \prod_{\substack{r|m \\ r \text{ prime}}} \left(1 + \frac{1}{r}\right) \leq \exp \sum_{\substack{r|m \\ r \text{ prime}}} \frac{1}{r} \leq \exp \sum_{\substack{r \leq m \\ r \text{ prime}}} \frac{1}{r} \leq c_2 \cdot \exp \log \log m.$$

This means that, for fixed p, the polynomial $\Phi_m(p, y) = 0$ has at most $c_2 \cdot m \log m$ roots (in y); in particular, each $p \in Q_T$ is isogenous to at most

$$\sum_{m=1}^{c_1 \cdot (\log T)^2} (c_2 \cdot m \log m) \leq c_3 \cdot (\log T)^4 \log \log T \leq c_3 \cdot (\log T)^5 \text{ points } q \in Q_T$$

for some explicit constant c_3. This implies the result. □

Remark 3.4.8. *The inequality*

$$\psi(m) \leq c_2 \cdot m \log m$$

which holds for every $m \geq 2$ with an absolute explicit constant c_2 will be often used throughout the document.

We now get rid of a special case for C, namely when C is the zero locus of some modular polynomial $\Phi_n(x, y) = 0$. If $(p, q) \in Q_T^2$ is isogenous to some point (x, y) of $C(\overline{\mathbb{Q}})$, then p and q are isogenous (i.e. p is isogenous to x, which is isogenous to y, which is isogenous to q).
For each $p \in Q_T$, the set

$$\{q \in Q_T \mid q \text{ is isogenous to } p\}$$

has size at most $c \cdot (\log T)^5$ by the above Lemma. This implies that on each vertical line of Q_T^2 only $c \cdot (\log T)^5$ points can be isogenous to some point of C and, adding all the vertical lines

of Q_T^2 together, we obtain that at most $2c \cdot T^2(\log T)^5$ points of Q_T^2 can be isogenous to some point of C.

Similarly, we get rid of when C is either a vertical line or an horizontal line. If C is the zero locus of the equation $x = x_0$, then the set

$$\{p \in Q_T \mid p \text{ is isogenous to } x\}$$

has size absolutely bounded by isogeny estimates (Theorem 3.2.2)

$$n' \leq c_0 \cdot h(x_0)^2 (2[\mathbb{Q}(x_0) : \mathbb{Q}]^2 \cdot \log^2 (2[\mathbb{Q}(x_0) : \mathbb{Q}])) \leq c_1$$

for a universal explicit constant c_0 and for some constant c_1 depending on x_0 only. Then we just need to remove finitely many lines and we obtain a bound of the shape $c_2 \cdot T^2$ for some constant c_2 independent of T. The proof for horizontal lines is identical.

We can therefore assume that C is not weakly special.

1. We construct a definable set for the application of the Pila-Wilkie Theorem 2.2.5. We fix $(p, q) \in Q_T^2$ and we consider the following set:

$$S_{p,q} = \{(a, b, c, d, a', b', c', d') \in \mathbb{R}^8 \mid \exists (x, y) \in C(\mathbb{C})$$
$$\frac{aj^{-1}(p) + b}{cj^{-1}(p) + d} = j^{-1}(x) \text{ and } \frac{a'j^{-1}(q) + b'}{c'j^{-1}(q) + d'} = j^{-1}(y)\}$$

   where j has domain $D \subseteq \mathbb{H}$, which is the standard fundamental domain for the action of the full modular group $\mathrm{SL}_2(\mathbb{Z})$, that we recall here:

$$D = \{z \in \mathbb{H} \mid -1/2 \leq \mathrm{Re}\, z < 1/2, |z| > 1\} \cup \{z \in \mathbb{H} \mid -1/2 \leq \mathrm{Re}\, z \leq 0, |z| = 1\}.$$

   Notice that if we vary p, q then $S_{p,q}$ varies definably. If $a, b, c, d, a', b', c', d'$ are integers, this means that (p, q) and (x, y) are isogenous.

2. Before we could use the Pila-Wilkie Theorem on $S_{p,q}$, we would need to understand which kind of semialgebraic varieties lie inside of $S_{p,q}$. In contrast to the application we gave before in section 2.3, here we actually have some positive dimensional semialgebraic part. Indeed, if we fix x and y, the values $a, b, c, d, a', b', c', d'$ are only determined by the linear conditions

$$aj^{-1}(p) + b = j^{-1}(x)(cj^{-1}(p) + d)$$
$$a'j^{-1}(q) + b' = j^{-1}(y)(c'j^{-1}(q) + d')$$

   which describe semialgebraic varieties contained in $S_{p,q}$. Notice that $a, b, c, d, a', b', c', d'$ determine x and y and we indeed will prove that any positive dimensional (connected) semialgebraic variety contained in $S_{p,q}$ corresponds to a fixed couple (x, y).

Suppose that we have a semialgebraic curve contained in S. Here the argument is fairly similar to that of 2.3. Shrinking to a suitable open subset of such curve, we parameterise $a, b, c, d, a', b', c', d'$ as algebraic functions

$$a(t), b(t), c(t), d(t), a'(t), b'(t), c'(t), d'(t)$$

of $t \in (0,1)$ (that, up to shrinking our interval, we can assume without branches). We argue that we can enlarge $(0,1)$ to a connected open subset $U \subseteq \mathbb{C}$. Indeed, we only need to check that the following holds true for some open neighborhood of t in U when it holds for a open neighborhood of t in $(0,1)$: the algebraic functions $a(t), b(t), c(t), d(t), a'(t), b'(t), c'(t), d'(t)$ determine uniquely the functions $x(t), y(t)$ (which parameterise x, y) and their analytic continuation on a (complex) neighborhood satisfies both

- $(x(t), y(t)) \in C(\mathbb{C})$ since $C(\mathbb{C})$ is (locally) the zero of holomorphic functions;
- $$\frac{a(t)j^{-1}(p) + b(t)}{c(t)j^{-1}(p) + d(t)} = j^{-1}(x(t))$$
  since the j-function is holomorphic (and similarly for a', b', c', d', y, q).

Again, as in section 2.3, one should take a slightly bigger open enlargement of D (e.g. because j^{-1} does not extend holomorphically on the image of the boundary in $Y(1)$); this is no real threat at all and for sake of simplicity we just forget about it.

Therefore, our $a(t), b(t), c(t), d(t), a'(t), b'(t), c'(t), d'(t)$ can be taken algebraic functions on a suitable open set $U \subseteq \mathbb{C}$ and therefore $j^{-1}(x(t))$ and $j^{-1}(y(t))$ are algebraic functions of t (recall that p and q are fixed).
Notice that $j^{-1}(x(t))$ and $j^{-1}(y(t))$ are mapped to points of $C(\mathbb{C})$ via the j-function: an (euclidean open subset of an) algebraic curve of $\mathbb{H}^2$ is mapped into an (euclidean open subset of an) algebraic curve of $Y(1)^2$ via j. This should remind us of the situation of section 2.3 and indeed this contradicts an Ax-Schanuel-type result for the j-function.
We use a result of Pila-Tsimerman [54]:

Theorem 3.4.9 (Pila-Tsimerman, Ax-Schanuel for the j-function)**.** *Let $u_1(t), \dots, u_n(t)$ be nonconstant holomorphic functions defined on a connected open subset $U \subseteq \mathbb{C}$, such that $u_i(t) \in \mathbb{H}$ for every index i and every $t \in U$. Suppose that for no positive integer m and indices $1 \leq p, q \leq n$ then $\Phi_m(j(u_p(t)), j(u_q(t))) = 0$ identically. Then the field*

$$\mathbb{C}(t, u_1(t), \dots, u_n(t), j(u_1(t)), \dots, j(u_n(t)))$$

has transcendence degree at least n over $\mathbb{C}(t)$.

Remark 3.4.10. *The principle beyond this statement is roughly: "all the algebraic (functional) relations satisfied by the j-function come from modular curves being the images (via j) of algebraic varieties".*

Here we set $n = 2$ with $u_1(t) = j^{-1}(x(t))$ and $u_2(t) = j^{-1}(y(t))$. Both are algebraic functions of t and moreover $(j(u_1(t)), j(u_2(t))) \in C(\mathbb{C})$ for every $t \in U$, i.e. the relevant transcendence degree is at most 1. We have two possibilities:

- For some positive integer n we have $\Phi_n(j(u_1(t)), j(u_2(t))) = 0$.
- Either $u_1(t)$ or $u_2(t)$ is constant.

Notice that we can actually *suppose* that not both $u_1(t)$ and $u_2(t)$ are constant: we already discussed semialgebraic curves corresponding to a fixed (x, y); if we require our (connected) semialgebraic curve to account for more than one point $(x, y) \in C(\mathbb{C})$, then at least one of the algebraic functions $j^{-1}(x(t))$ (i.e. $u_1(t)$) and $j^{-1}(y(t))$ (i.e. $u_2(t)$) needs to be nonconstant.
The two possibilities above then force respectively C to be special (i.e. the zero locus of a modular polynomial $\Phi_n(x, y) = 0$) or C to be either a vertical or an horizontal line (we are implicitly using that two distinct irreducible algebraic curves intersect at finitely many points). We dealt with both these exceptional cases before.

Therefore, we can conclude that $S_{p,q}$ contains no positive dimensional (connected) semialgebraic variety accounting for more than one $(x, y) \in C(\mathbb{C})$.

3. We can then apply the Pila-Wilkie Theorem 2.2.5 (with a constant independent of (p, q) by uniform definability); notice that semialgebraic connected curves (and hence all positive dimensional connected semialgebraic sets) only account for fixed (x, y). We obtain:

   For every $\varepsilon > 0$ there is a constant e (independent of (p, q)) such that any set of rational points of $S_{p,q}$ (each point corresponding to a different $(x, y) \in C(\mathbb{C})$) with exponential height bounded by H has size at most $e \cdot H^{\varepsilon}$ (for any real $H \geq 1$).

4. We now combine isogeny estimates and the outcome of the Pila-Wilkie Theorem above in order to obtain an upper bound on the degrees of the connecting isogenies, in terms of T. Indeed, suppose that $(p, q) \in Q_T$ and $(x, y) \in C(\overline{\mathbb{Q}})$ are isogenous and the minimal isogeny has a "big" degree; then also $[\mathbb{Q}(x, y) : \mathbb{Q}]$ is "big" and this observation allows us to produce "too many" rational points with bounded height on $S_{p,q}$.
   Let $d = \max(2, [\mathbb{Q}(x, y) : \mathbb{Q}])$. By isogeny estimates (Theorem 3.2.2), we have positive integers m, n such that:

$$\Phi_m(p, x) = 0, \Phi_n(q, y) = 0 \text{ and } m, n \leq c_5 \cdot d^2 (\log d)^2 (\log T)^2$$

   for some universal explicit constant c_5.
   The condition $\Phi_m(p, x) = 0$ produces an integral point on $S_{p,q}$:

$$\frac{aj^{-1}(p) + b}{cj^{-1}(p) + d} = j^{-1}(x)$$

for some integers a, b, c, d such that $\gcd(a, b, c, d) = 1$ and $ad - bc = m$; moreover, since $j^{-1}(p)$ and $j^{-1}(x)$ are confined to D, we have, by Lemma 2.5.12,

$$\max(|a|, |b|, |c|, |d|) \leq m.$$

The analogous inequality holds for a', b', c', d' exploiting the condition $\Phi_n(q, y) = 0$ and we can conclude:

$(p, q) \in Q_T$ and $(x, y) \in C(\overline{\mathbb{Q}})$ being isogenous produces a rational point of $S_{p,q}$ (corresponding to (x, y)) of exponential height at most $c_5 \cdot d^2(\log d)^2(\log T)^2$.

We now use a Galois-theoretic argument in order to produce many of these points (and corresponding to distinct couples (x, y)).
Let K be the (number) field of definition of C (see Remark 3.4.3 above). Each conjugate of the couple (x, y) over K lies on C and these are at least $d/[K : \mathbb{Q}]$ many. Since the modular polynomials $\Phi_k(\cdot, \cdot)$, for $k \geq 1$, have integer coefficients, one also has $\Phi_m(p, x') = 0$ and $\Phi_n(q, y') = 0$ for any conjugate (x', y') of (x, y), i.e. (p, q) is isogenous to (x', y').
Therefore, each conjugate gives rise to integers $(a, b, c, d, a', b', c', d')$ which satisfy the same bound as above in terms of the height. Moreover, distinct pairs of conjugates give different integers (since $(a, b, c, d, a', b', c', d')$ determine (x, y)) corresponding to distinct values of (x, y), so we obtain:

there is a set of rational points of S corresponding to distinct pairs (x, y) with exponential height bounded by $c_5 \cdot d^2(\log d)^2(\log T)^2$ and size $\geq d/[K : \mathbb{Q}]$.

We compare this with the outcome of 3. setting $\varepsilon = 1/3$ and $H = c_5 \cdot d^2(\log d)^2(\log T)^2$, obtaining the inequality

$$d/[K : \mathbb{Q}] \leq e \cdot (c_5 \cdot d^2(\log d)^2(\log T)^2)^{1/3}$$

for positive constants e and c_5 (notice that e is ineffective). Since $[K : \mathbb{Q}]$ is fixed, one obtains

$$d/(\log d)^2 \leq c_6 \cdot (\log T)^2$$

for some positive constant c_6 and thus

$$d \leq c_7 \cdot (\log T)^3$$

for some positive constant c_7. From the inequality given by isogeny estimates above we also conlude

$$m, n \leq c_8 \cdot (\log T)^9$$

for some positive constant c_8, i.e. we have bounded the degree of the connecting isogenies in terms of T.

5. We conclude the argument by showing that the "isogenous images" of $C(\mathbb{C})$ via these isogenies of bounded degree are very far from covering Q_T^2 (and in this sense "most" points of Q_T^2 will not be isogenous to any point of $C(\mathbb{C})$).
Let us define our isogenous images of C.

$$C_{m,n} = \{(x,y) \in Y(1)^2 \mid \exists (x',y') \in C$$

$$\Phi_m(x,x') = 0 \text{ and } \Phi_n(y,y') = 0\}.$$

Namely, $C_{m,n}(\mathbb{C})$ is the set of points that are isogenous to some point of $C(\mathbb{C})$ with prescribed isogenies (of degree m and n respectively); the union, over all the positive integers m, n, of the curves $C_{m,n}(\mathbb{C})$ is precisely the set of points of $Y(1)^2$ isogenous to some point of $C(\mathbb{C})$.
Notice that $C_{m,n}$ is closed (as it follows from observing that, for a fixed m, if $\Phi_m(x,x') = 0$ and x tends to infinity then x' does too); we do not need this in the proof.

Suppose that C is defined by a polynomial of degree at most F in each variable. We claim that $C_{m,n}$ is defined by a polynomial of degree at most $8\psi(m)\psi(n)F$ in each variable; this is implicit in [42], Lemma 3.1, where the result is proven by taking the resultants corresponding to the elimination of x' and y'.
One can also observe that $C_{m,n}$ is the projection (that forgets the last two coordinates) of the curve

$$C'_{m,n} = \{(x,y,x',y') \in Y(1)^4 \mid \Phi_m(x,x') = 0, \Phi_n(y,y') = 0 \text{ and } (x',y') \in C\}$$

and $C'_{m,n}$ is the intersection of three hypersurfaces of degree $2\psi(m), 2\psi(n)$ and $2F$ respectively. Here the relevant degree for hypersurfaces is the sum of the degrees of a (minimal) defining polynomial in each coordinate. We refer to section 3.7 for a deeper discussion of this specific degree, that is more suitable for powers of $Y(1)$ than the usual "algebro-geometric" degree.
The claim follows taking the product of these bounds and observing that coordinate projections do not increase the degree.

We have that $(p,q) \in Q_T^2$ is isogenous to a point of $C(\mathbb{C})$ with $\Phi_m(p,x) = 0$ and $\Phi_n(q,y) = 0$ if and only if $(p,q) \in C_{m,n}(\mathbb{C})$. We deduced at the end of 4. that

$$m, n \leq c_8 \cdot (\log T)^9$$

and

$$\psi(m) \leq c_2 \cdot m \log m \text{ and } \psi(n) \leq c_2 \cdot n \log n.$$

These inequalities combined with the observation above imply that $C_{m,n}$ is defined by a polynomial of degree (i.e. the sum of the degrees in each variable) bounded by

$$16F\psi(m)\psi(n) \leq 16Fc_2^2 \cdot (c_8 \cdot (\log T)^9)^2 \cdot \log^2 (c_8 \cdot (\log T)^9) \leq c_9 \cdot (\log T)^{19}$$

for some positive constant c_9. This implies that the union

$$Z = \bigcup_{m,n} C_{m,n} \text{ for } m, n \leq c_8 \cdot (\log T)^9$$

is the zero locus of a (reducible) polynomial of degree at most $c_{10} \cdot (\log T)^{37}$ for some positive constant c_{10}.

Every $(p, q) \in Q_T^2$ which is isogenous to some point of $C(\mathbb{C})$ lies in Z. We observe that every vertical line intersects Z in at most $c_{10} \cdot (\log T)^{37}$ points, unless some irreducible component of some $C_{m,n}$ is itself a vertical line: this possibility implies indeed that C is a vertical line itself. This is in fact a consequence of this more general observation: if W is an irreducible component of $C_{m,n}$, then $C \subseteq W_{m,n}$. Alternatively, one can observe that Z contains at most $c_{10} \cdot (\log T)^{37}$ vertical lines (by degree considerations).
Summing over all the vertical lines of Q_T^2, we conclude that the number of points of Q_T^2 which are isogenous to some point of $C(\mathbb{C})$ is at most

$$c_{10} \cdot 2T^2 (\log T)^{37}$$

and this concludes the proof of Proposition 3.4.1.

Remark 3.4.11. *A careful analysis of the proof above shows that our "avoiding points" can also be found on a fixed curve C' provided $C' \not\subseteq C_{m,n}$ for any positive integers m, n (this condition is of course necessary).*
For instance, suppose that C' is a rational line. The proof above is exactly the same for the parts 1.-4., from which we deduce $m, n \leq c_8 \cdot (\log T)^9$ whenever a point $(p, q) \in Q_T^2$ is isogenous to a point of $C(\mathbb{C})$ with isogenies of degree m, n.
For the last part 5., the points of $C'(\mathbb{Q}) \cap Q_T^2$ which are isogenous to some point of $C(\mathbb{C})$ are contained in the union of the curves $C_{m,n}$, for $m, n \leq c_8 (\log T)^9$, which is a (reducible) algebraic curve Z of degree bounded by $c_{10}(\log T)^{37}$. Hence, the points of $Z \cap C'$ are at most $c_{10}(\log T)^{37}$ (here we are using that $C' \not\subseteq Z$, i.e. $C \not\subseteq C_{m,n}$ for any m, n).
Now it is enough to observe that the set $C'(\mathbb{Q}) \cap Q_T^2$ has size bounded below by $c' \cdot T^2$ for some positive constant c'.
In general C' might not have enough rational points (e.g. if it has genus greater or equal to 2 it has only finitely many), but we can always find plenty of suitably "small" degree.

6. Let now assume C given by the unit equation

$$C = \{(x, y) \in Y(1)^2 \mid y = x + 1\}.$$

The proof up to 4. is exactly the same; we borrow the terminology for the isogenous image $C_{m,n}$ of C from 5. above.
We assume the following fact, whose proof is postponed to 7.

$$\text{If } C \subseteq C_{m,n} \text{ then } m = n = 1.$$

Suppose $(p,q) \in Q_T$ is isogenous to a point $(x,y) \in C(\overline{\mathbb{Q}})$, with connecting isogenies of degree m and n:

$$\Phi_m(x,p) = 0 \text{ and } \Phi_n(y,q) = 0.$$

We recall from 4. that

$$m, n \leq c_8 \cdot (\log T)^9$$

for some positive constant c_8. We split into two possibilities:

- If $(x,y) \in C(\mathbb{Q})$, then the following result of Kenku [33] provides an explicit constant K_0 such that

  $$m, n \leq K_0$$

  i.e. two isogenous elliptic curves over $\mathbb{Q}$ are connected by an isogeny of degree bounded by K_0. Thus, we construct Z as

  $$Z = \bigcup_{m,n} C_{m,n} \text{ for } m, n \leq K_0$$

  and observe that if $(p,q) \in Q_T^2 \setminus Z$ then (p,q) cannot be isogenous to any point $(x,y) \in C(\mathbb{Q})$.

- Suppose now that $(x,y) \in C(\overline{\mathbb{Q}})$ and $[\mathbb{Q}(x,y) : \mathbb{Q}] \geq 2$. Since C is defined over $\mathbb{Q}$ and the modular polynomials $\Phi_k(\cdot,\cdot)$ have rational coefficients, there must be a Galois conjugate (x',y') of (x,y) over $\mathbb{Q}$ such that

  $$\Phi_m(x',p) = 0, \Phi_n(y',q) = 0 \text{ and } (x',y') \in C(\overline{\mathbb{Q}}).$$

  We now construct an isogeny connecting (x',y') to (x,y). Indeed, there are integers M, N such that

  $$\Phi_M(x',x) = 0 \text{ and } \Phi_N(y',y) = 0 \text{ with } M \leq m^2 \text{ and } N \leq n^2.$$

  This can be easily seen on lattices: if $\Lambda_1 \subseteq \Lambda_2 \subseteq \Lambda_3$ is an inclusion of lattices, then $[\Lambda_3 : \Lambda_1] = [\Lambda_3 : \Lambda_2] \cdot [\Lambda_2 : \Lambda_1]$ (but the degree of the connecting *cyclic* isogeny can strictly decrease, since the quotient Λ_3/Λ_1 needs not to be cyclic whenever the intermediate quotients Λ_3/Λ_2 and Λ_2/Λ_1 are).
  Notice that not both M, N will be equal to 1, since by assumption on the degree of the field of definition we have $(x,y) \neq (x',y')$.
  By definition of $C_{M,N}$, this means that

  $$(x,y) \in C \cap C_{M,N}.$$

  We can bound M, N with

  $$c_8^2 \cdot (\log T)^{18}$$

for the positive constant c_8 above, which is independent of T, and we wish to estimate the size of the set

$$\bigcup_{M,N} (C \cap C_{M,N}) \text{ with } M, N \leq c_8^2 \cdot (\log T)^{18}, \text{ not both equal to 1.}$$

Observe that the degree (in the sense of 5., i.e. the sum of the degrees in each coordinate of a minimal defining polynomial) of $C_{M,N}$ is bounded above by $16\psi(M)\psi(N)$, since in our case C has degree 2.
We also use the estimate (from Remark 3.4.8, that follows from Mertens Theorem) that we proved before:

$$\psi(m) \leq c_2 \cdot m \log m \text{ and } \psi(n) \leq c_2 \cdot n \log n$$

for some explicit constant c_2. Then the size above can be estimated by the upper bound

$$\sum_{M,N \leq c_8^2 \cdot (\log T)^{18}} (32\psi(M)\psi(N)) \leq \sum_{M,N \leq c_8^2 \cdot (\log T)^{18}} (32c_2^2 \cdot M \log M \cdot N \log N) \leq c_{11} \cdot (\log T)^{73}$$

for some constant c_{11} indepedent of T.

Therefore, any $(p, q) \in Q_T^2 \setminus Z$ is isogenous to a point in a set of bounded size ($\leq c_{11} \cdot (\log T)^{73}$) via isogenies of bounded degree ($\leq c_8^2 \cdot (\log T)^{18}$). Using again the estimates following from Mertens Theorem, we see that the possibilities for (p, q) are at most

$$(c_{11} \cdot (\log T)^{73}) \cdot \sum_{M,N \leq c_8^2 \cdot (\log T)^{18}} (\psi(M)\psi(N)) \leq c_{13} \cdot (\log T)^{146}$$

for some constant c_{13} independent of T. This concludes the proof.

7. We finally prove:

$$\text{If } C \subseteq C_{m,n} \text{ then } m = n = 1$$

when C is given by the unit equation

$$C = \{(x, y) \in Y(1)^2 \mid y = x + 1\}.$$

Let $(t, t+1) \in C(\mathbb{C})$, where we think of t as a very big positive real number. Since $C \subseteq C_{m,n}$, there must be some $(u, u+1) \in C(\mathbb{C})$ such that

$$\Phi_m(t, u) = 0 \text{ and } \Phi_n(t+1, u+1) = 0.$$

We take preimages in the fundamental domain D: let $\alpha, \beta, \alpha', \beta' \in D$ such that

$$j(\alpha) = t, j(\beta) = u, j(\alpha') = t+1, j(\beta') = u+1$$

and notice, as a consequence of Lemma 2.5.12, that there are integers a, b, c, d and a', b', c', d' with

$$\gcd(a, b, c, d) = 1 \text{ and } \gcd(a', b', c', d') = 1$$

$$ad - bc = m \text{ and } a'd' - b'c' = n$$

$$\max(|a|, |b|, |c|, |d|) \leq m \text{ and } \max(|a'|, |b'|, |c'|, |d'|) \leq n$$

such that

$$\frac{a\alpha + b}{c\alpha + d} = \beta \text{ and } \frac{a'\alpha' + b'}{c'\alpha' + d'} = \beta'.$$

Notice that a, b, c, d can attain finitely many values (for fixed m, n) and as t tends to infinity then $\alpha = j^{-1}(t)$ tends to infinity (in D) as well so that

$$\frac{a\alpha + b}{c\alpha + d} \text{ tends to } \frac{a}{c}$$

unless $c = 0$. Since rational numbers are not on the boundary of D we conlude that $c = 0$ as long as t is big enough. The same holds for $t + 1$ tending to infinity and we obtain $c' = 0$. Observe that these imply $a/d > 0$ and $a'/d' > 0$ (by $ad - bc = m$ and $a'd' - b'c' = n$).

We now exploit the q-expansion of the j-function, namely

$$j(\tau) = q^{-1} + 744 + O(|q|) \text{ for } \tau \in D$$

where $q = \exp(2\pi i\tau)$. Let $v = \exp(2\pi i\alpha)$ and $w = \exp(2\pi i\alpha')$. We have

$$j(\alpha) = v^{-1} + 744 + O(|v|) \text{ and } j(\alpha') = w^{-1} + 744 + O(|w|)$$

and comparing with $j(\alpha') = j(\alpha) + 1$ we obtain (notice that $|v|, |w|$ tend to zero as t tends to infinity)

$$w^{-1} = v^{-1} + 1 + O(|v|).$$

We also compute

$$j(\beta) = j\left(\frac{a\alpha + b}{d}\right) = \mu \cdot v^{-a/d} + 744 + O\left(\left|v^{a/d}\right|\right)$$

for $\mu = \exp(-2\pi i(b/d))$ and

$$j(\beta') = j\left(\frac{a'\alpha + b'}{d'}\right) = \mu' \cdot w^{-a'/d'} + 744 + O\left(\left|w^{a'/d'}\right|\right)$$

for $\mu' = \exp(-2\pi i(b'/d'))$. Notice that in general $v^{-a/d}$ (and $w^{-a'/d'}$) is a "multi-valued" expression. However, in this particular scenario, v is a small positive real number (since $\alpha = j^{-1}(t)$ lies on the imaginary axis); here μ is accounting for the correct determination. We can finally compare $j(\beta') = j(\beta) + 1$:

$$\mu' \cdot w^{-a'/d'} = \mu \cdot v^{-a/d} + 1 + O\left(|v|^{1/(m+n)}\right)$$

where $O\left(|v|^{1/(m+n)}\right)$ is obtained from $m+n$ being an upper bound for

$$\max\left(|a|,|b|,|c|,|d|,|a'|,|b'|,|c'|,|d'|\right).$$

We substitute the estimate for w^{-1} above in the last equation:

$$\mu' \cdot v^{-a'/d'}\left(1+\frac{a'}{d'}\cdot v+O(|v|^2)\right)=\mu\cdot v^{-a/d}+1+O\left(|v|^{1/(m+n)}\right).$$

Recall that the integers a,b,c,d,a',b',c',d' can only assume finitely many values and therefore there are finitely many possibilities for μ and μ' as well. Then, as t tends to (plus) infinity then v^{-1} becomes "very big", from which we compare the leading terms above and deduce $a/d=a'/d'$ and $\mu=\mu'$.
Substituting into the above we can write

$$\mu\cdot\frac{a}{d}\cdot v^{(1-a/d)}\left(1+O(|v|)\right)=1+O\left(|v|^{1/(m+n)}\right)$$

from which, again comparing the leading terms, we deduce $a/d=1$ and $\mu=1$, i.e. $b=0$. Unwinding the definitions, this implies that $m=n=1$.

3.5 Points avoiding rational lines

In this section, we focus on a result of a different shape than anything we discussed above: before, we found a point "avoiding" a single curve, while now we wish to find a point which "avoids" countably many different varieties.
We also introduce a useful result by Pazuki (which is a modern version of a classical result by Faltings [22]), which states that the heights of two isogenous points are comparable (in terms of the degree of the connecting isogeny).

For a number field K we find it convenient, as before, to introduce the sets

$$K_T=\{x\in Y(1)(K)\mid h(x)<\log T\}.$$

We note that the size of K_T is roughly T^{2d} (as a consequence of the estimates by Schanuel in [59]).

Proposition 3.5.1. *Let K be a number field of degree $d\geq 11$.*

1. *For "most" $p\in Y(1)(K)$, we have that p is not isogenous to any $x\in Y(1)(\mathbb{Q})$. More precisely, there is a universal positive constant c' such that at most*

$$c'\cdot T^2(d\cdot\log T)^{52}$$

many elements of K_T are isogenous to any $x\in Y(1)(\mathbb{Q})$.

2. *For "most" $(p,q)\in Y(1)^2(K)$, we have that (p,q) is not isogenous to any $(x,y)\in Y(1)^2(\overline{\mathbb{Q}})$ which lies on a rational line; namely, any $(x,y)\in Y(1)^2(\overline{\mathbb{Q}})$ for which there are $a,b\in\mathbb{Q}$ such that $y=ax+b$.*

Remark 3.5.2. *We need 1. in order to prove 2.*

Remark 3.5.3. *Of course one must restrict the degree of the field of definition of the lines that should be avoided: any point on $Y(1)^2(\overline{\mathbb{Q}})$ is on some line over a number field.*
Viceversa, K needs to have some "big" field of definition: any point of $Y(1)^2(\mathbb{Q})$ (and even any point of $Y(1)^2(\overline{\mathbb{Q}})$ whose field of definition has degree 2 over $\mathbb{Q}$) lies on a rational line.

We discuss generalised versions of this result in the section 3.6.

We introduce the following result by Pazuki (based on a previous result of Faltings [22]), which roughly says that the j-invariants of two isogenous elliptic curves have comparable height in terms of the degree of the connecting isogeny. This is the main result of [48].

Theorem 3.5.4 (Pazuki)**.** *Let $x, y \in Y(1)(\overline{\mathbb{Q}})$ and let $n \geq 2$ be a positive integer such that $\Phi_n(x, y) = 0$. Then:*

$$|h(x) - h(y)| \leq 12 \log n.$$

We now prove Proposition 3.5.1.

1. The elements of K_T have height bounded above by $\log T$. Observe that if $p \in K_T$ is isogenous to some $x \in \mathbb{Q}$, then there is a connecting isogeny with degree bounded above by

   $$c \cdot (d \cdot \log T)^{2.1}$$

   for some universal positive constant c, by isogeny estimates (Theorem 3.2.2). By the result of Pazuki (Theorem 3.5.4), we obtain

   $$h(x) \leq \log T + 26 \log \log T + 26 \log d + 12 \log c$$

   and therefore x can assume at most $2c^{24} \cdot T^2 (d \cdot \log T)^{52}$ values (since x is rational). Now we can bound the number of admissible values of p, since we know a bound on the degree of the connecting isogeny and at most $\psi(n)$ values of p are connected to x via an isogeny of degree n; we obtain that the quantity of such values is bounded above by

   $$2c^{24} \cdot T^2 (d \cdot \log T)^{52} \cdot \sum_{n=1}^{c \cdot (d \cdot \log T)^{2.1}} \psi(n) < 2c^{27} c_0^2 \cdot T^2 (d \cdot \log T)^{57}$$

   where we are using that for any positive integer n we have $\psi(n) \leq c_0 \cdot n \log n$ for some universal positive constant c_0. These are far from being all the roughly T^{2d} many elements of K_T and we are done.

2. We give a proof of 2. Let us take $(p, q) \in K_T^2$ and let ℓ be a line in $Y(1)^2$. We define informally $S_{\ell,p,q}$ as

   "the set consisting of the entries of the matrices that connect (p, q) to a point on $\ell(\mathbb{C})$"

so that integral points of $S_{\ell,p,q}$ represent an isogeny connecting (p,q) to some point on ℓ. The family $S_{\ell,p,q}$ varies definably in ℓ, p, q (since ℓ varies among all the lines, which are varieties of bounded algebro-geometric degree). We wish to use the Pila-Wilkie Theorem 2.2.5 in order to get a bound on the degrees of the connecting isogenies. Let $(x,y) \in Y(1)^2(\overline{\mathbb{Q}})$ be a point isogenous to (p,q) (say x is isogenous to p and y is isogenous to q). We can use isogeny estimates (Theorem 3.2.2) and we obtain the upper bound on the degree of each connecting isogeny

$$c \cdot (d \cdot [\mathbb{Q}(x,y) : \mathbb{Q}] \cdot \log T)^{2.1}$$

for c a universal positive constant (notice that $h(p)$ and $h(q)$ are bounded above by $\log T$).

Let us fix a (non weakly special) rational line over which (x,y) lies; exactly as before, we obtain an upper bound on the degree of $\mathbb{Q}(x,y)$ over $\mathbb{Q}$ of the shape

$$[\mathbb{Q}(x,y) : \mathbb{Q}] \leq e \cdot d \cdot \log T$$

for a constant e independent of T (but depending on d). Moreover, e is independent of our rational line (and our point (p,q)), since the Pila-Wilkie Theorem is applied uniformly in the family of all rational lines (together with an extra "base" point in $Y(1)^2$). For reference, this is exactly the possibility 2. of the proof seen in the previous section.

We needed ℓ non weakly special (namely, neither vertical, nor horizontal nor equal to $Y_1 = \{(x,y) \in Y(1)^2 \mid x = y\}$ - these are precisely the exceptions that would arise from the Ax-Schanuel Theorem for the j-function, see Theorem 3.4.9) because for weakly special subvarieties the Pila-Wilkie Theorem 2.2.5 fails, at least to the extent that we are aiming to use it for now.

We now assume that (p,q) is not isogenous to any element of $Y(1)^2(\mathbb{Q})$ (this can be done safely, thanks to 1., by removing a small portion of K_T^2) and let $(x,y) \in Y(1)^2(\overline{\mathbb{Q}})$ be isogenous to (p,q). Suppose that $(x,y) \in Y(1)^2(\overline{\mathbb{Q}})$ lies on a rational line and let $a, b \in \mathbb{Q}$ such that $y = ax + b$. Since (x,y) is not a rational point it will have at least one Galois conjugate over $\mathbb{Q}$, say (x', y'), lying on the very same rational line, i.e. satisfying $y' = ax' + b$. Of course a line is determined by two points:

$$a = \frac{y' - y}{x' - x}, \ b = y - ax.$$

Using elementary height properties (namely, those of Proposition 2.6.7), we have

$$h(a) \leq h(x) + h(x') + h(y) + h(y') + 5;$$

$$h(b) \leq h(x) + h(y) + h(a) + 2.$$

We can, as in 1. above, find a bound on the heights $h(x), h(y)$ (and hence on $h(x'), h(y')$ by the invariance of the Weil height under Galois conjugation) due to (x,y) being isogenous

to a point of K_T^2, using isogeny estimates (that we used above already) and the result of Pazuki. We can combine the bound obtained by isogeny estimates above together with the outcome of the Pila-Wilkie Theorem and we bound from above the degree of a connecting isogeny with

$$c' \cdot (d \cdot \log T)^{4.2}$$

for a constant c' that doesn't depend on the rational line. We can then apply the result of Pazuki (Theorem 3.5.4) and we obtain

$$h(x), h(y), h(x'), h(y') \leq \log T + 52 \log\log T + 52 \log d + 12 \log c'$$

and hence (up to enlarging c a bit)

$$h(a) \leq 4 \log T + 208(\log\log T + \log d + \log c);$$

$$h(b) \leq 6 \log T + 312(\log\log T + \log d + \log c).$$

We conclude using these observations, where we assume that d is fixed and T is big enough:

(a) There are at most T^{21} choices for the rational line containing (x, y), for T big, as $h(a) < 4.1 \log T$ and $h(b) < 6.1 \log T$;

(b) (p, q) is contained in one of the isogenous images of these lines through isogenies of degree $\leq (\log T)^5$ as long as T is big enough;

(c) (p, q) is contained in a finite union of at most $T^{21}(\log T)^{10}$ (possibly reducible) curves of algebro-geometric degree at most $(\log T)^{30}$, hence in a (probably reducible) curve E of degree $T^{21}(\log T)^{40}$. We are using that the isogenous image of a curve of degree F via isogenies of degree m and n has degree bounded above by $8F\psi(m)\psi(n)$.

But this last condition holds sparsely in K_T^2, as $d \geq 11$ (notice that we had not used such condition yet); for instance, a vertical line intersects E in at most $T^{21}(\log T)^{40}$ points, while it contains roughly T^{2d} many (a vertical line can be contained in E, but E can contain at most $T^{21}(\log T)^{40}$ vertical lines). This proves 2.

Remark 3.5.5. *During the proof we restricted to the case of K being a number field of degree $d \geq 11$, but the same argument holds if we replace K by the set $\mathcal{S}$ of all algebraic numbers of degree $\leq d$ over $\mathbb{Q}$ (and in this case any $d \geq 5$ would do). One would have*

$$\mathcal{S}_T = \{x \in Y(1)(\overline{\mathbb{Q}}) \mid h(x) < \log T \text{ and } [\mathbb{Q}(x) : \mathbb{Q}] \leq d\}$$

and in this case the size of $\mathcal{S}_T$ is roughly $T^{d(d+1)}$ by the main result of the paper by Masser-Vaaler [40].

3.6 Some general principles of avoiding problems

Let us state some general principles underlying our "avoiding problems", guided by the examples we have discussed above. Our aim is to extend our strategy for finding "avoiding points" to the furthest extent.

We can formulate the principle behind the Masser-Zannier Theorem in a more abstract way.

Given a Shimura variety S and a proper subvariety $X \subseteq S$, then, for d big enough, "most" points of $S(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not "isogenous" to any point of $X(\overline{\mathbb{Q}})$.

We do not pause on what "isogenous" means for an arbitrary Shimura variety; for us S will always be either a power of the multiplicative group $\mathbb{G}_{\mathrm{m}}$, an abelian variety, a power of $Y(1)$ or some $\mathcal{A}_g$. We wish instead to generalise this statement allowing X to vary, as we have already done taking into account all the rational lines.

Given a Shimura variety S and a countable family of proper subvarieties $X_1, X_2, \ldots \subseteq S$ of "diophantine geometric origin", then, for d big enough, "most" points of $S(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not "isogenous" to any point of $X(\overline{\mathbb{Q}})$.

We give an example. Let us take $S = Y(1)^2$ and we take the X_i to be the rational lines. As we have seen in the previous section, we have the following.

Proposition. *Let $d \geq 11$. Then for "most" points $(p, q) \in Y(1)^2(\overline{\mathbb{Q}})$ (of degree $\leq d$) there are no $m, n \geq 1$, $(x, y) \in Y(1)^2(\overline{\mathbb{Q}})$ and $a, b \in \mathbb{Q}$ such that:*

$$\Phi_m(p, x) = \Phi_n(q, y) = 0 \text{ and } y = ax + b.$$

Remark 3.6.1. *We point out that, rather than required our "avoiding points" to lie in a fixed number field, we can, with the same proof, obtain the results above only fixing their algebraic degree.*

This Proposition can be pushed further, in a few directions:

- we can increase the dimension of the ambient space $Y(1)^r$;
- we can ask for our lines to be defined on "small" extensions of $\mathbb{Q}$;
- instead of a family of lines we can take arbitrary varieties of bounded "algebro-geometric degree".

We clarify what we mean by "algebro-geometric degree" in the next section. We give a first statement.

Theorem 3.6.2. *Let a, b, c, k be positive integers and let $\mathcal{F}$ be a set consisting of subvarieties of $Y(1)^r$ of dimension $\leq a$ and "algebro-geometric" degree $\leq b$, each defined over a number field of degree $\leq c$ over $\mathbb{Q}$. Then, for d big enough, "most" points $P \in Y(1)^r(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any $Q \in X(\overline{\mathbb{Q}})$ for any $X \in \mathcal{F}$.*

Remark 3.6.3. *We observe that the conditions on the "algebro-geometric" degree and on the degree of the field of definition are necessary. Indeed, every point of $Y(1)^r(\overline{\mathbb{Q}})$ is contained in both:*

- *a variety of small "algebro-geometric" degree and unbounded degree of its field of definition (e.g. the point itself);*
- *a variety defined over $\mathbb{Q}$ but with unbounded "algebro-geometric" degree: any point $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in Y(1)^r(\overline{\mathbb{Q}})$ is contained in the hypersurface defined by $P(x_1) = 0$ where $P(t)$ is the minimal polynomial of α_1 over $\mathbb{Q}$.*

Remark 3.6.4. *Notice that each individual member of $\mathcal{F}$ is defined over a number field of degree bounded by c, but the compositum of such number fields can be an infinite extension of $\mathbb{Q}$. For instance, the family of lines*

$$\ell_t = \{(x, y) \in Y(1)^2 \mid y = x + t\} \text{ for } t \text{ such that } [\mathbb{Q}(t) : \mathbb{Q}] = 2$$

is an admissible choice for $\mathcal{F}$.

The scheme of the proof is not conceptually very different from the one we gave for $\mathcal{F}$ consisting of lines over $\mathbb{Q}$. There are two main obstacles:

- we had used before that a line is determined by two points and hence we could recover the height of the line from the height of the points. In higher dimensions there is a complication: a plane, for instance, might not be determined by any finite number of points, if they all lie on a line. However, it turns out that the height of this very line is bounded in terms of the points... This is essentially a technicality and one can get rid of this issue using an inductive argument;
- a more serious issue is related to the presence of some weakly special subvariety satisfying $W \subseteq X$ for some $X \in \mathcal{F}$. In the case of a family of lines we have dealt with these issues using ad hoc arguments: we can classify precisely the lines that contain positive dimensional weakly special subvarieties (which are precisely the vertical and horizontal lines and the unique "diagonal" modular curve). The matter in higher dimension becomes much subtler and its solution requires substantially different arguments.
  The obstruction comes from the fact that the Pila-Wilkie Theorem 2.2.5 "fails" whenever X contains a positive dimensional weakly special W: our definable set can contain semi-algebraic varieties accounting for more than a point of X. We solve the issue proving a generalised form of isogeny estimates, that reads roughly:

  There is a polynomial $P(t_1, t_2, t_3) \in \mathbb{Q}[t_1, t_2, t_3]$, depending on the ambient space $Y(1)^r$ only, with the following property. If a point p is isogenous to some point $q \in W$, then there is a positive integer n such that

  $$n \leq P(h(p), \deg W, [\mathbb{Q}(W, p) : \mathbb{Q}])$$

  and such that p is isogenous, via an isogeny of degree n, to a point $q' \in W$.

The Pila-Wilkie Theorem is now used to bound $[\mathbb{Q}(W):\mathbb{Q}]$ instead of $[\mathbb{Q}(q):\mathbb{Q}]$ as we did before.

We can also extend the realm of our "avoiding problems" in another direction, namely restricting our source of "most" points to a subvariety of S.

Given a Shimura variety S, a proper subvariety $X \subseteq S$ and a subvariety $Y \subseteq S$ which is not "related" to X, then, for d big enough, "most" points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not "isogenous" to any point of $X(\overline{\mathbb{Q}})$.

We give an example. Let $C_1, C_2 \subseteq Y(1)^2$ be two curves. We wish to find (many) points on C_1 that are not isogenous to any point on C_2. The following is an obstruction:

"There exist m, n such that for any $(p_1, q_1) \in C_1(\mathbb{C})$ there is $(p_2, q_2) \in C_2(\mathbb{C})$ with $\Phi_m(p_1, p_2) = 0$ and $\Phi_n(q_1, q_2) = 0$.".

This is actually the only obstruction for us; in this case we say that C_1 and C_2 are "isogenous curves". Referring to the notation introduced in the section 3.4, this is the same as $C_1 \subseteq (C_2)_{m,n}$.

We can reinterpret this condition in terms of special subvarieties of the product. We consider $C_1 \times C_2 \subseteq Y(1)^2 \times Y(1)^2$, with coordinates (x_1, y_1, x_2, y_2) coming from the projections onto the four factors. The condition above is the same as $C_1 \times C_2$ intersecting a two-dimensional special subvariety of the form "Zero locus of $\Phi_m(x_1, x_2)$ and $\Phi_n(y_1, y_2)$" in a one-dimensional variety (which is "unlikely", as the intersection should be zero-dimensional).

Proposition 3.6.5. *Let $C_1, C_2 \subseteq Y(1)^2$ be two non-isogenous curves defined over $\overline{\mathbb{Q}}$. There is a positive integer d_0 such that, for any $d \geq d_0$, "most" points of $C_1(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $C_2(\overline{\mathbb{Q}})$.*

In terms of the reasoning above with $C_1 \times C_2 \subseteq Y(1)^4$, this Proposition means that the projection of

$$\bigcup_{m,n \geq 1} (C_1 \times C_2) \cap (\{(x_1, y_1, x_2, y_2) \mid \Phi_m(x_1, x_2) = 0 \text{ and } \Phi_n(y_1, y_2) = 0\})$$

on C_1 is "sparse" in $C_1(\overline{\mathbb{Q}})$.

Remark 3.6.6. *It can be proven that the projection above is sparse also allowing all the two-dimensional special subvarieties in the union, i.e.*

$$\bigcup_{\substack{Y \text{ special} \\ \dim Y = 2}} (C_1 \times C_2) \cap Y.$$

This is a consequence of the following argument. The intersection of a curve $C \subseteq Y(1)^2$ with the union of all the modular curves Y_n is very sparse in C by isogeny estimates:

$$\text{If } (x, y) \in Y_n \text{ then } n \leq c \cdot (h(x, y) \cdot [\mathbb{Q}(x, y) : \mathbb{Q}])^{2.1}$$

for a universal explicit constant c, whenever (x, y) is not a special point (and only finitely many special points can lie on C, as a consequence of the proven André-Oort Conjecture). If we bound above the degree of (x, y), then the height $h(x, y)$ grows like an algebraic function of n, while the number of points with height bounded by T on $C(\overline{\mathbb{Q}})$ grows like an exponential of T, at least when $[\mathbb{Q}(x, y) : \mathbb{Q}]$ is big enough (but fixed); we therefore produce many "avoiding" points.
The "exceptional" points of bounded degree are even conjecturally finitely many, by SGO (Conjecture 3.2).

We finally mix all the settings: we both allow X to vary in a "geometric family" and we fix some subvariety $Y \subseteq S$. For a variety $X \subseteq Y(1)^r$, we define:

$$X^{(i)} \text{ is the subvariety of } Y(1)^n \text{ consisting of the points } p \text{ such that } p \text{ is isogenous to some point of } X \text{ via a connecting isogeny of degree } i.$$

This is the final statement for $Y(1)^r$.

Theorem 3.6.7. *Let a, b, c, r be positive integers and let $\mathcal{F} = \{X_1, X_2, \ldots\}$ be a set of subvarieties of $Y(1)^r$ of dimension $\leq a$, "algebro-geometric" degree $\leq b$ and field of definition of degree $\leq c$ over $\mathbb{Q}$. Let $Y \subseteq Y(1)^r$ be an irreducible subvariety such that for any positive integers i and n we have $Y \cap X_i^{(n)}$ not Zariski dense in Y. Then, for d big enough, "most" points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $X_i(\overline{\mathbb{Q}})$ for any $X_i \in \mathcal{F}$.*

3.7 Geometric degrees

In the following few sections we prove Theorem 3.6.7, our main generalisation of the result by Masser and Zannier.
We will extensively use an "algebro-geometric" notion of degree throughout the document and we clarify it here.

Let n be a positive integer and let X be an irreducible subvariety of $Y(1)^n$ of dimension d, defined over $\mathbb{C}$. We take coordinates $(x_1, x_2, \ldots, x_n)$ on $Y(1)^n$ corresponding to the projections onto its n factors. For each choice of d positive integers $1 \leq i_1 < i_2 < \ldots < i_d \leq n$ and complex numbers $\underline{c} = (c_1, c_2, \ldots, c_d) \in Y(1)^d(\mathbb{C})$ the linear subaspace $\mathcal{L}_{\underline{c}}$ defined by

$$\begin{cases} x_{i_1} = c_1 \\ x_{i_2} = c_2 \\ \ldots \\ x_{i_d} = c_d \end{cases}$$

has codimension d in $Y(1)^n$ and "usually" the intersection will consist of a finite set of points of fixed cardinality; namely, there is a non-empty open subset U of $Y(1)^d$ such that if $\underline{c} = (c_1, c_2, \ldots, c_d) \in U$ then

$$X \cap \mathcal{L}_{\underline{c}} \text{ consists precisely of } N(i_1, i_2, \ldots, i_d) \text{ points}$$

where the constant $N(i_1, i_2, \ldots, i_d)$ is independent of $\underline{c}$ (as long as $\underline{c} \in U$). Notice that $N(i_1, i_2, \ldots, i_d)$ can be possibly equal to zero, for instance if X is contained in some $\mathcal{L}_{\underline{c}}$.
Alternatively, we can construct the projection

$$\pi : Y(1)^n \to Y(1)^d$$

$$\pi(x_1, x_2, \ldots, x_n) = (x_{i_1}, x_{i_2}, \ldots, x_{i_d})$$

and we observe that $N(i_1, i_2, \ldots, i_d)$ is just the "typical number of preimages" of π; the quantity $N(i_1, i_2, \ldots, i_d)$ is zero exactly whenever π, restricted to X, is not a dominant map (and one can always find a choice of $i_1, i_2, \ldots, i_d$ such that π is dominant, taking algebraically independent coordinates $x_{i_1}, x_{i_2}, \ldots, x_{i_d}$ over X).

Definition. *For an irreducible variety $X \subseteq Y(1)^n$ of dimension d, we define its degree* $\deg X$ *as the sum, over all the possible choices of positive integers $1 \leq i_1 < i_2 < \ldots < i_d \leq n$, of the quantities*

$$N(i_1, i_2, \ldots, i_d).$$

Whenever X is reducible, we set its degree as the sum of the degrees of its irreducible components.

Example 3.7.1. *The degree of a subset $S \subseteq Y(1)^n$ consisting of k points is k.*

Example 3.7.2. *The degree of a hypersurface $S \subseteq Y(1)^n$ defined by a polynomial $P(x_1, x_2, \ldots, x_n)$ is the sum of the individual degrees* $\deg_{x_1} P + \deg_{x_2} P + \ldots + \deg_{x_n} P$.

Remark 3.7.3. *For our purposes, there is no difference between the degree we just defined and the more common degree given by the embedding of $Y(1)^n \cong \mathbb{A}^n \subseteq \mathbb{P}_n$ in the usual way. Both just give a description of the "geometric complexity" of the relevant variety and they are not intrinsic properties of such variety (in the sense that they also depend on the ambient space, e.g. they depend on n in this case).*
We prefer to stick to our choice of degree, which corresponds to the "obvious" embedding of $Y(1)^n$ into $(\mathbb{P}_1)^n$.

The usual notion of degree is defined as follows. Let X be an irreducible subvariety of $\mathbb{P}_n$ of dimension d. The family of linear subspaces $\mathcal{L}_u \subseteq \mathbb{P}_n$ of codimension d is parameterised by a grassmanian, say G, which is an algebraic variety. Again, there is a non-empty open subset $U \subseteq G$ such that

$$\text{if } u \in U \text{ then } X \cap \mathcal{L}_u \text{ consists precisely of } N \text{ points}$$

where N is independent of u (as long as $u \in U$). The value N is the classical algebro-geometric degree of X (together with an embedding into $\mathbb{P}_n$) and we denote it with $\deg_c X$.

Essentially, for our purposes $\deg X$ and $\deg_c X$ are equivalent. We prefer our unconventional

choice of $\deg X$ since it corresponds to the embedding of $Y(1)^n$ into $(\mathbb{P}_1)^n$ (together with an appropriate intersection form).

The next Proposition shows that $\deg X$ and $\deg_c X$ are comparable and we mostly use it to deduce a few useful properties from classical Intersection Theory.

Proposition 3.7.4. *Let $X \subseteq Y(1)^n$ be a subvariety of dimension d. Then*

$$\deg_c X \leq \deg X \leq \binom{n}{d} \deg_c X.$$

Proof. Let us fix a projection

$$\pi : Y(1)^n \to Y(1)^d$$

$$\pi(x_1, x_2, \ldots, x_n) = (x_{i_1}, x_{i_2}, \ldots, x_{i_d})$$

such that π is dominant when restricted to X and let $N(i_1, i_2, \ldots, i_d)$ the cardinality of a generic fiber of π. We can observe that

$$N(i_1, i_2, \ldots, i_d) \leq \deg_c X$$

by definition of $\deg_c X$, since any linear space of codimension d arising as fiber of π cannot intersect X in more than $\deg_c X$ (yet the generic fiber intersects X into $N(i_1, i_2, \ldots, i_d)$ points). Since for the choices of $i_1, i_2, \ldots, i_d$ for which π is not dominant when restricted to X we have set $N(i_1, i_2, \ldots, i_d) = 0$, one sums up all the inequalities above and obtains, by definition of $\deg X$,

$$\deg X \leq \binom{n}{d} \deg_c X.$$

□

Remark 3.7.5. *These inequalities are sharp. The upper inequality is an equality in the generic case, e.g. whenever X is a linear subspace with the property that no x_i is constant on X. On the opposite extreme, the lower inequality becomes an equality whenever X is a linear subspace of dimension d with the property that exactly d coordinates x_i are nonconstant on X; another example is the hypersurface given by the equation*

$$x_1 \cdot x_2 \cdot \ldots \cdot x_n = 1.$$

We essentially will only use the following two properties of the degree:

Proposition 3.7.6. *Let $X, Y \subseteq Y(1)^n$ be subvarieties and let Z be an irreducible component of $X \cap Y$. Then*

$$\deg Z \leq 2^n \cdot \deg X \cdot \deg Y.$$

Proof. This is a consequence of the classical property of the usual degree, for instance Theorem 7.7 of the classic Hartshorne book [31],

$$\deg_c Z \leq \deg_c X \cdot \deg_c Y$$

together with Proposition 3.7.4 that allows us to compare the two notions of degree. □

Proposition 3.7.7. *The family $\mathcal{S}$ consisting of the subvarieties of $Y(1)^n$, defined over $\mathbb{C}$, of a given dimension and a given degree, is definable over $\mathbb{C}$ (in the language of rings).*

Proof. The subvarieties of $\mathbb{P}_n$ of a given dimension and a given classical degree are parameterised by a variety (namely a Chow Variety, see [24] for reference). The strategy is observing that each of the conditions that are implied in the definition of our degree become definable inside the Chow Variety.
We quote the following result.

Let $\varphi_s : X_s \to Y_s$ be a definable family of morphisms parameterised by S. The subsets of S consisting of:

- $s \in S$ such that φ_s is dominant;
- $s \in S$ such that φ_s is dominant and the generic fiber is finite of cardinality N.

are all definable (for any positive integer N). Moreover, there is some positive constant C such that, for every $s \in S$ such that φ_s is dominant and the generic fiber is finite, such generic fiber has cardinality bounded above by C.

The quoted result together with the existence of the (definable) Chow Variety for the classical degree guarantee that the subset consisting of points with a prescribed behaviour (e.g. having dominant projection or a given generic fiber cardinality) for π, for any choice of $1 = i_1 < i_2 < \ldots < i_d = n$, is definable. Combining all the finitely many possibilities, the result follows. □

Our construction of a definable set that parameterises varieties of a given degree could actually give us more information. We need to put some height on the parameters that define our X_i, as we did already with rational lines, namely ordering the rational lines by the height of their coefficients.

Indeed, for the rational line $\{(x, y) \mid y = ax + b\}$ the height of the pair (a, b) seems an appropriate measure of its complexity. We can perform the same construction more in general. As the subvarieties of $Y(1)^r$ of bounded degree are parameterised by a variety $\mathcal{S}$ with the nice property that

$$V \text{ is defined over } K \text{ if and only if it corresponds to a point of } \mathcal{S}(K)$$

as it can be seen by an immediate Galois-theoretic argument: the set of varieties of a given degree is invariant by Galois-conjugation, hence $\mathcal{S}$ must be defined over $\mathbb{Q}$.
Given a variety V defined over $\overline{\mathbb{Q}}$, we take the height $h(V)$ of V as the height of the point representing V in $\mathcal{S}(\overline{\mathbb{Q}})$.

We conclude the section with a technical Lemma which will be useful later.

Lemma 3.7.8. *Let $\mathcal{S}$ be a variety over $\overline{\mathbb{Q}}$ parameterising some subvarieties of $Y(1)^r$ of a given degree and a given dimension. Let k be a positive integer. There exists a constant c', depending on $\mathcal{S}$ and k only, satisfying the following property.*
Given points $p_1, p_2, \ldots, p_k \in Y(1)^r(\overline{\mathbb{Q}})$ such that there exists a unique $V \in \mathcal{S}$ containing them all, then

$$h(V) < c' \cdot (1 + h(p_1) + h(p_2) + \ldots + h(p_k))$$

where $h(V)$ is the height of the point representing V in $\mathcal{S}$.

Remark 3.7.9. *This is essentially an interpolation statement: a variety uniquely determined by some points cannot be "too much more complicated" than such points. An example of this principle is the Lagrange method for interpolating polynomials.*

Proof. The proof uses abstract functorial properties of the height (as the explicit interpolation argument would be too messy). The argument is a direct consequence of definability and point 1. (which is the "obvious" one) of Theorem 2.7.2.

Let $Z \subseteq (Y(1)^r)^k(\mathbb{C})$ consisting of the k-uples of points $(p_1, p_2, \ldots, p_k)$ that determine uniquely some $V \in \mathcal{S}$ containing them all. This is a definable (over $\mathbb{C}$, in the language of rings) condition, essentially amounting to the existence and uniqueness of a point of the variety $\mathcal{S}$ for which some equations vanish at $(p_1, p_2, \ldots, p_k)$.
As Z is a definable set (over $\mathbb{C}$, in the language of rings), it is a constructible set and therefore the map that associates to a point of Z (that consists of k-tuple of points) the unique point of $\mathcal{S}$ representing a variety containing the k-tuple is a morphism φ from Z to $\mathcal{S}$, provided we split Z into finitely many normal subvarieties. One has that φ is actually defined over some number field K over which $\mathcal{S}$ is defined: if the points $p_1, p_2, \ldots, p_k$ are all defined over K, then the unique variety parameterised by $\mathcal{S}$ containing them must be defined over K (otherwise any Galois-conjugate would satisfy the same property, contradicting uniqueness).
One can then use by 1. Theorem 2.7.2 that for any $z \in Z$

$$h(\varphi(z)) < c' \cdot (1 + h(z))$$

for some positive c' depending on Z only (and hence on $\mathcal{S}$ and k only). Observing that $\varphi(p_1, p_2, \ldots, p_k)$ is the point representing V in $\mathcal{S}$ we conclude the proof.

□

Remark 3.7.10. *One needs to work over $\mathbb{C}$ in order to use that a 1-to-1 correspondence between normal varieties is a morphism: for instance, the cubic root*

$$t \to t^{1/3}$$

is a 1-to-1 correspondence over $\mathbb{R}$ which is not a morphism.

3.8 Classification of weakly special subvarieties

In this section we give an explicit description of the weakly special subvarieties of $Y(1)^r$.

We denote with $x_1, x_2, \ldots, x_r$ the coordinates on $Y(1)^r$ arising from the projection onto its r factors.

Definition. *A weakly special subvariety of $Y(1)^r$ is an irreducible component of a system of equations either of the shape*

$$\Phi_n(x_i, x_j) = 0$$

for any positive integer n and integers $1 \leq i, j \leq r$ and

$$x_i = \gamma$$

for some $1 \leq i \leq r$ and some complex number γ.

Remark 3.8.1. *As in the case of special subvarieties the condition $\Phi_n(x_i, x_j) = 0$, whenever $i = j$, amounts to a finite union of conditions of the shape $x_i = s$ where s is a singular modulus.*

Remark 3.8.2. *The only difference between weakly special subvarieties and special subvarieties is that, in the former case, γ is not required to be a singular modulus, but it can be any complex number.*

Remark 3.8.3. *Each point of $Y(1)^r$ is a weakly special point. Moreover, all the "vertical" linear subspaces, namely subspaces defined by a system of the shape*

$$\begin{cases} x_{i_1} = \gamma_1 \\ x_{i_2} = \gamma_2 \\ \ldots \\ x_{i_d} = \gamma_d \end{cases}$$

for some integers $1 \leq i_1 < i_2 < \ldots < i_d \leq r$ and complex numbers $\gamma_1, \gamma_2, \ldots, \gamma_d$ are weakly special subvariety. Every point is contained in some positive dimensional weakly special subvariety.

We now proceed to give a rather accurate description of weakly special subvarieties. We need a few Lemmas.

Lemma 3.8.4. *Let m and n be distinct positive integers. The set of solutions to the system*

$$\begin{cases} \Phi_m(x_1, x_2) = 0 \\ \Phi_n(x_1, x_2) = 0 \end{cases}$$

is finite and it consists of special points of $Y(1)^2$.

Proof. This amounts to observe that the modular curves $Y_m, Y_n \subseteq Y(1)^2$ are distinct irreducible curves, hence their intersection is a finite union of points, which are special by definition (they arise as irreducible components of a system of the correct shape). □

Lemma 3.8.5. *Let m and n be positive integers. The set of solutions to the system*

$$\begin{cases} \Phi_m(x_1, x_2) = 0 \\ \Phi_n(x_2, x_3) = 0 \end{cases}$$

is contained in the union of the varieties defined by

$$\Phi_d(x_1, x_3) = 0$$

over the positive integers d such that $d \mid mn$ and mn/d is a perfect square.

Proof. This is directly seen on lattices. Suppose that for elliptic curves E_1, E_2, E_3 corresponding to lattices $\Lambda_1, \Lambda_2, \Lambda_3 \subseteq \mathbb{C}$ we have that

$$E_1 \text{ and } E_2 \text{ are related by a cyclic isogeny of degree } m;$$
$$E_2 \text{ and } E_3 \text{ are related by a cyclic isogeny of degree } n.$$

Rescaling the lattices by appropriate homotheties, we obtain that $\Lambda_1 \subseteq \Lambda_2 \subseteq \Lambda_3$, where the quotients Λ_2/Λ_1 and Λ_3/Λ_2 are cyclic of order m and n respectively. Thus, the quotient Λ_3/Λ_1 has order mn, although it might not be cyclic. However, there exists an appropriate positive integer e such that e^2 divides mn and $\Lambda_1 \subseteq \frac{1}{e}\Lambda_3$, so that

$$E_1 \text{ and } E_3 \text{ are related by a cyclic isogeny of degree } mn/e^2.$$

□

Let us now consider a system of modular equations of the shape

$$\begin{cases} \Phi_{n_{1,2}}(x_1, x_2) = 0 \\ \Phi_{n_{1,3}}(x_1, x_3) = 0 \\ \dots \\ \Phi_{n_{1,r}}(x_1, x_r) = 0 \\ \Phi_{n_{2,3}}(x_2, x_3) = 0 \\ \dots \\ \Phi_{n_{r-1,r}}(x_{r-1}, x_r) = 0 \end{cases}$$

for positive integers $n_{i,j}$ (whenever $1 \leq i < j \leq n$) defining a subvariety of $Y(1)^r$. We call such a system *full*, in the sense that each pair of coordinates x_i and x_j for any $1 \leq i < j \leq r$ are related by a modular equation $\Phi_{n_{i,j}}(x_i, x_j) = 0$.
We argue that all the irreducible components of a full system are either points or curves: indeed, there cannot be two algebraically independent coordinates x_i and x_j (for any choice of $1 \leq i < j \leq r$).
Notice that these irreducible components are not only weakly special, but even special.

Lemma 3.8.6. *Let $C \subseteq Y(1)^r$ be a weakly special curve with the property that, for each $1 \leq i \leq r$, the projection*

$$x_i : C \to Y(1)$$

is dominant. Then C is an irreducible component of a full system as above. In particular, C is special.

We say that a curve of such a shape is *dominant special curve*. We state our characterisation first and we prove both the Lemma and this at once.

Proposition 3.8.7. *Let $W \subseteq Y(1)^r$ be a weakly special subvariety of dimension d. Up to permutation of coordinates, there is a decomposition*

$$Y(1)^{r_1} \times Y(1)^{r_2} \times \ldots \times Y(1)^{r_d} \times Y(1)^{r'} = Y(1)^r$$

with dominant special curves

$$C_i \subseteq Y(1)^{r_i}$$

for each $1 \leq i \leq d$ and a point $(\gamma_1, \gamma_2, \ldots, \gamma_{r'}) \in Y(1)^{r'}$ such that

$$W = C_1 \times C_2 \times \ldots \times C_d \times \{(\gamma_1, \gamma_2, \ldots, \gamma_{r'})\}.$$

Proof. We start by permuting coordinates and supposing that

$$x_i : W \to Y(1)$$

is not dominant precisely in the last r' coordinates. Setting, $\gamma_i = x_{i+r'}(w)$ for any point $w \in W$ (that gives rise to a unique value since W is irreducible) we obtain that W is of the form

$$W' \times \{(\gamma_1, \gamma_2, \ldots, \gamma_{r'})\}$$

and that x_i is dominant, when restricted to W, whenever $i \leq r - r'$. Notice that W' is a weakly special subvariety, as we can check that substituting values γ_i instead of the coordinate x_i in the definition of weakly special subvarieties still gives a weakly special subvariety.
Therefore, we can assume that $W' \subseteq Y(1)^{r-r'}$ is a weakly special subvariety such that all the projections

$$x_i : W' \to Y(1)$$

for $1 \leq i \leq r - r'$ are dominant. Let us then assume in the statement such property for W, i.e. that all the projections onto factors are dominant, and forget about W' and r'.

Since W (or C) is weakly special, it is defined as a zero locus of a system of the shape

$$\Big\{ \Phi_{n_t}(x_{i_t}, x_{j_t}) = 0$$

for a positive integer T and positive integers $n_1, n_2, \ldots, n_T$ and $1 \leq i_1, i_2, \ldots, i_T, j_1, j_2, \ldots, j_T \leq r$, together with some extra condition of the shape

$$\Big\{ x_i = \gamma$$

that we can safely forget since x_i is dominant when restricted to W (or C); this already proves that W (or C) is special.
By Lemma 3.8.4, there are no pairs (i_t, j_t) with $i_t = j_t$ and moreover no pairs (i_{t_1}, j_{t_1}) and

(i_{t_2}, j_{t_2}) which are equal (i.e. $i_{t_1} = i_{t_2}$ and $j_{t_1} = j_{t_2}$, or $i_{t_1} = j_{t_2}$ and $j_{t_1} = i_{t_2}$). Thus, we can rewrite our system as

$$\left\{\Phi_{n_{i,j}}(x_i, x_j) = 0\right.$$

for choices of (i, j) in a subset S of

$$\{(p, q) \in \mathbb{N}^2 \mid 1 \leq p < q \leq r\}.$$

Let us informally say that $1 \leq i < j \leq r$ are "connected" if there are a positive integer k and pairs $(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k) \in S$ such that $i = i_1, j = j_k$. We argue that "connected indices form a full system", in the sense that by repeatedly applying Lemma 3.8.5 and possibly permuting coordinates we obtain that W (or C) is an irreducible component of

$$Z_1 \times Z_2 \times \ldots \times Z_d \subseteq Y(1)^{r_1} \times Y(1)^{r_2} \times \ldots \times Y(1)^{r_d} = Y(1)^r$$

where each Z_i is an irreducible component of the zero locus of a full system for $Y(1)^{r_i}$, for some choice of positive integers $r_1, r_2, \ldots, r_d$ such that $r_1 + r_2 + \ldots + r_d = r$. As we have seen above the irreducible components of zero loci of full systems are either curves or points; the latter case is not admissible since the projection

$$Y(1)^r \to Y(1)^{r_i}$$

cannot have image consisting of just a point when restricted to W (or C) for any choice of $1 \leq i \leq d$, since it factors through some dominant projection

$$Y(1)^r \to Y(1)^{r_i} \to Y(1)$$

for instance choosing the coordinate x_k with $k = r_1 + r_2 + \ldots + r_{i-1} + 1$. This means that W (or C) is precisely a product of d dominant special curves; $d = 1$ in the case of C. □

Remark 3.8.8. *The converse is not true: full systems might define points only. For instance, the zero locus of the system*

$$\begin{cases} \Phi_2(x_1, x_2) = 0 \\ \Phi_3(x_1, x_3) = 0 \\ \Phi_5(x_2, x_3) = 0 \end{cases}$$

is contained in the zero locus of

$$\begin{cases} \Phi_2(x_1, x_2) = 0 \\ \Phi_6(x_2, x_3) = 0 \\ \Phi_5(x_2, x_3) = 0 \end{cases}$$

and this situation implies, by Lemma 3.8.5, that such a set consists only of finitely many special points.

Even the condition $n_{i,k} \mid n_{i,j} \cdot n_{j,k}$ with $n_{i,j} \cdot n_{j,k}/n_{i,k}$ for any choice of $1 \leq i, j, k \leq r$ with i, j, k

distinct positive integers is not sufficient. The system

$$\begin{cases} \Phi_2(x_1, x_2) = 0 \\ \Phi_2(x_1, x_3) = 0 \\ \Phi_2(x_1, x_4) = 0 \\ \Phi_2(x_1, x_5) = 0 \\ \Phi_4(x_2, x_3) = 0 \\ \Phi_4(x_2, x_4) = 0 \\ \Phi_4(x_2, x_5) = 0 \\ \Phi_4(x_3, x_4) = 0 \\ \Phi_4(x_3, x_5) = 0 \\ \Phi_4(x_4, x_5) = 0 \end{cases}$$

defines only special points. Indeed, for any value of x_1, two among x_2, x_3, x_4, x_5 must be equal (since $\psi(2) = 3$) and thus an additional condition

$$\Phi_1(x_i, x_j) = 0$$

for some $2 \leq i < j \leq 5$ appears.
It is not clear to us whether a simple combinatorial condition for full systems to have a positive dimensional zero locus exists, besides looking at a "compatible" choice of matrices representing the isogenies in $\mathbb{H}$; one can just look at upper triangular matrices (e.g. at the behaviour at infinity).

Let us discuss a notion of geometric complexity for weakly special subvarieties of $Y(1)^r$. Let us start with a full system

$$\begin{cases} \Phi_{n_{1,2}}(x_1, x_2) = 0 \\ \Phi_{n_{1,3}}(x_1, x_3) = 0 \\ \dots \\ \Phi_{n_{1,r}}(x_1, x_r) = 0 \\ \Phi_{n_{2,3}}(x_2, x_3) = 0 \\ \dots \\ \Phi_{n_{r-1,r}}(x_{r-1}, x_r) = 0 \end{cases}$$

such that C is a positive dimensional irreducible component of its zero locus. Notice that C has to be a curve.

Definition. *We set the special degree of C as*

$$\deg_s C = \max_{1 \leq i < j \leq r} (n_{i,j}).$$

Whenever W is a weakly special subvariety which is the product of the dominant special curves $C_1, C_2, \dots, C_d$ and a point, we set the complexity of W as

$$\deg_s W = \deg_s C_1 \cdot \deg_s C_2 \cdot \ldots \cdot \deg_s C_d.$$

As one could expect from the notation, the special degree is comparable with the usual algebro-geometric degree and with the degree we defined above.

This is has a nice consequence for us: an upper bound on the algebro-geometric degree of the zero locus of

$$\Phi_n(x_1, x_2) = 0$$

is therefore an upper bound on the degree n of the connecting isogeny.

Let us formally prove it.

Proposition 3.8.9. *Given a positive integer r, there are positive constants c, e (depending on r only) satisfying the following.*
Let $W \subseteq Y(1)^r$ be a weakly special subvariety. Then

$$c \cdot (\deg_c W)^e \leq \deg_s W \leq \deg_c W$$

where $\deg_c$ denotes the classical algebro-geometric degree given by the embedding $Y(1)^r \cong \mathbb{A}^r \subseteq \mathbb{P}_r$.

Remark 3.8.10. *In view of Proposition 3.7.4, one also obtain the same bound (up to changing c) for the degree in $Y(1)^r$:*

$$c \cdot (\deg W)^e \leq \deg_s W \leq \deg W.$$

Proof. Since the classical degree of a point is 1 and for any affine varieties X, Y

$$\deg_c X \times Y = \deg_c X \cdot \deg_c Y$$

we just need to prove our statement in the case of W a special dominant curve C.
Let C be then an irreducible curve arising as an irreducible component of the full system

$$\begin{cases} \Phi_{n_{1,2}}(x_1, x_2) = 0 \\ \Phi_{n_{1,3}}(x_1, x_3) = 0 \\ \dots \\ \Phi_{n_{1,r}}(x_1, x_r) = 0 \\ \Phi_{n_{2,3}}(x_2, x_3) = 0 \\ \dots \\ \Phi_{n_{r-1,r}}(x_{r-1}, x_r) = 0 \end{cases}$$

and let us suppose that $n = n_{1,2}$ is maximal among the $n_{i,j}$, for $1 \leq i < j \leq r$ (this can be obtained by permutation of coordinates). The projection π

$$\pi : Y(1)^r \to Y(1)^2$$

$$\pi(x_1, x_2, \dots, x_r) = (x_1, x_2)$$

satisfies $\pi(C)$ equal to the zero locus of

$$\Phi_n(x_1, x_2) = 0$$

(which is the modular curve $Y_n \subseteq Y(1)^2$).

As a property of the classical degree through projection, we obtain the following inequality

$$\deg_c(Y_n) \leq \deg_c(C)$$

and we observe that

$$\psi(n) \leq \deg_c(Y_n) \leq 2\psi(n)$$

since $\Phi_n(x_1, x_2)$ is a polynomial of degree $\psi(n)$ in each variable. Clearly

$$n \leq \psi(n) \leq \deg_c(C)$$

but n is just $\deg_s(C)$, so we obtain the upper bound.

For the lower bound, we notice that each hypersurface $H_{i,j}$ defined as the zero locus of

$$\Phi_{n_{i,j}}(x_i, x_j) = 0$$

for any $1 \leq i < j \leq r$ has classical degree bounded above by $2\psi(n_{i,j})$, as $\Phi_{n_{i,j}}(x_i, x_j)$ is a polynomial of degree $\psi(n_{i,j})$ in each coordinate. Since C is an irreducible component of the intersection of all these hypersurfaces, we obtain:

$$\deg_c C \leq \deg_c \left(\bigcap_{1 \leq i < j \leq r} H_{i,j} \right) \leq \prod_{1 \leq i < j \leq r} \deg_c H_{i,j} \leq \prod_{1 \leq i < j \leq r} 2\psi(n_{i,j})$$

since, for the classical degree, the degree of an intersection is bounded above by the product of the degrees of the varieties involved. Recall that there is a positive constant C such that for every positive integer m

$$\psi(m) < C \cdot m \log m$$

as a consequence of Merten's Theorem (Remark 3.4.8). We can therefore bound

$$2\psi(n_{i,j}) < 2C \cdot n_{i,j}^2 \leq 2C \cdot n^2$$

for every $1 \leq i < j \leq r$ and, recalling that $n = \deg_s C$, we obtain the desired lower bound.

□

We will observe later that the geometric degree of the maximal weakly special subvarieties of a given variety is bounded above in terms of the degree of such variety, by a result of Binyamini-Daw [9].

We have seen above that every $W \subseteq Y(1)^r$ can be written (up to permutation of the coordinates) as a product of dominant special curves and a point:

$$W = C_1 \times C_2 \times \ldots \times C_d \times \{(\gamma_1, \gamma_2, \ldots, \gamma_{r-r'})\}.$$

Definition. *We say that $C_1 \times C_2 \times \ldots \times C_d$ is the special type of W.*

Remark 3.8.11. *We want to keep track of the permutation of the coordinates in this case: for us the zero locus of*

$$\Phi_n(x_1, x_2) = 0$$

and that of

$$\Phi_n(x_1, x_3) = 0$$

are not of the same special type. Essentially, two weakly special subvarieties have the same special type exactly when their "nonconstant part" is the same.

Lemma 3.8.12.

- *There are only finitely many special types of weakly special subvarieties $W \subseteq Y(1)^r$ satisfying an upper bound on the special degree (or, equivalently, the classical degree).*
- *Let $W_1, W_2 \subseteq Y(1)^r$ two weakly special subvarieties of the same special type. If $W_1 \neq W_2$, then W_1 and W_2 are disjoint.*

We conclude the section by proving an analogue of isogeny estimates which takes into account weakly special subvarieties.

We find it convenient to introduce a complexity notion on isogenies between two points $p, q \in Y(1)^r(\overline{\mathbb{Q}})$ as well. If $p = (p_1, p_2, \ldots, p_r)$ and $q = (q_1, q_2, \ldots, q_r)$ are related by an isogeny so that

$$\Phi_{n_1}(p_1, q_1) = 0, \Phi_{n_2}(p_2, q_2) = 0, \ldots, \Phi_{n_r}(p_r, q_r) = 0$$

for appropriate positive integers $n_1, n_2, \ldots, n_r$, we say that p and q are related by an isogeny of *complexity* n, where $n = n_1 \cdot n_2 \cdot \ldots \cdot n_r$.

Remark 3.8.13. *It is in general not true that every isogeny between points p and q has to give an isogeny from an individual p_i to q_i (even permuting coordinates) due to the presence of "mixed" isogenies, that can arise when two among the p_i are isogenous (the correct analogue should be an $r \times r$ matrix collecting the isogenies from each p_{i_1} to each q_{i_2}). It is however true (by semisimplicity of abelian varieties) that if the points p and q are isogenous then they must be related by an isogeny of the shape above, since the simple factors must be pairwise isogenous.*

Proposition 3.8.14. *Given a positive integer r, there exists a polynomial $P(t_1, t_2, t_3, t_4) \in \mathbb{Q}[t_1, t_2, t_3, t_4]$ with the following property.*
Let $W \subseteq Y(1)^r$ be a weakly special subvariety defined over $\overline{\mathbb{Q}}$ and let $p \in Y(1)^r(\overline{\mathbb{Q}})$. Suppose that p is isogenous to some point of $W(\overline{\mathbb{Q}})$; then there exists a positive integer n satisfying

$$n < P(h(p), [\mathbb{Q}(p), \mathbb{Q}], \deg W, [\mathbb{Q}(W) : \mathbb{Q}])$$

such that p is isogenous to a point of $W(\overline{\mathbb{Q}})$ via an isogeny of complexity n.

Proof. This essentially boils down to applying isogeny estimates with our explicit description of W.

Let us write $p = (p_1, p_2, \ldots, p_r)$ and let us suppose W is of the form

$$W = C_1 \times C_2 \times \ldots \times C_d \times \{(\gamma_1, \gamma_2, \ldots, \gamma_{r-r'})\}$$

for dominant special curves $C_1, C_2, \ldots, C_d$.

1. Let us prove the result in the case of W equal to a dominant special curve C which is an irreducible component of the full system

$$\begin{cases} \Phi_{n_{1,2}}(x_1, x_2) = 0 \\ \Phi_{n_{1,3}}(x_1, x_3) = 0 \\ \ldots \\ \Phi_{n_{1,r}}(x_1, x_r) = 0 \\ \Phi_{n_{2,3}}(x_2, x_3) = 0 \\ \ldots \\ \Phi_{n_{r-1,r}}(x_{r-1}, x_r) = 0 \end{cases}$$

and let $N = \max_{i,j}(n_{i,j})$. We consider the projection on the first coordinate $C \to Y(1)$, which is surjective; this means that there is a point $q = (q_1, q_2, \ldots, q_r) \in C$ such that $q_1 = p_1$. We argue that p and q are connected by an isogeny of "small" complexity.

First, notice that since p is isogenous to a point of C, this means that all the coordinates p_i and p_j (for $1 \leq i, j \leq r$) represent isogenous elliptic curves. Thus, for any $1 \leq i \leq r$, we have that p_i is isogenous to q_i. Moreover, since

$$\Phi_{n_{1,i}}(p_1, q_i) = \Phi_{n_{1,i}}(q_1, q_i) = 0$$

we deduce that

$$[\mathbb{Q}(p_1, q_i) : \mathbb{Q}(p_1)] \leq \psi(n_{1,i}) < C \cdot N^2$$

for a universal constant C as a consequence of Merten's Theorem (Remark 3.4.8). We can use isogeny estimates Theorem 3.2.2 and we obtain that, for some explicit constant c, there is some positive integer m_i such that

$$\Phi_{m_i}(p_i, q_i) = 0$$

$$m_i < c \cdot ([\mathbb{Q}(p_i, q_i) : \mathbb{Q}]h(p_i))^{2.1}$$

and by

$$[\mathbb{Q}(p_i, q_i) : \mathbb{Q}] \leq [\mathbb{Q}(p_1, p_i, q_i) : \mathbb{Q}(p_1, q_i)][\mathbb{Q}(p_1, q_i) : \mathbb{Q}(p_1)][\mathbb{Q}(p_1) : \mathbb{Q}] < C \cdot [\mathbb{Q}(p) : \mathbb{Q}]^2 N^2$$

we obtain that, for some universal constant C' we have

$$m_i < C' \cdot [\mathbb{Q}(p) : \mathbb{Q}]^{4.2} \cdot h(p)^{2.1} \cdot N^2.$$

Choosing all the appropriate $m_1, m_2, \ldots, m_r$, we construct an isogeny from p to q of complexity bounded by a polynomial in $[\mathbb{Q}(p) : \mathbb{Q}], h(p)$ and $\deg C$.

2. By the first part, we can choose points on $C_1, C_2, \ldots, C_d$ whose coordinates are isogenous to the corresponding coordinates of p and these choices can be made using a complexity bounded by the correct polynomial.
   We just need to take care of $(\gamma_1, \gamma_2, \ldots, \gamma_{r-r'})$. This is simply the usual isogeny estimates: since by hypothesis $p_{r'+i}$ is isogenous to γ_i for every $1 \leq i \leq r - r'$, we can choose a connecting isogeny of degree bounded by

$$c \cdot h(p_{r'+i})^{2.1} \cdot [\mathbb{Q}(p_{r'+i}, \gamma_i) : \mathbb{Q}]^{2.1} \leq c \cdot h(p)^{2.1} \cdot [\mathbb{Q}(p) : \mathbb{Q}]^{2.1} \cdot [\mathbb{Q}(W) : \mathbb{Q}]^{2.1}$$

   for a universal explicit constant c. Multiplying all these bounds together we obtain an isogeny of degree bounded above by an appropriate polynomial in $h(p), [\mathbb{Q}(p) : \mathbb{Q}]$ and $[\mathbb{Q}(W) : \mathbb{Q}]$.
   Recall that the bounds arising from the C_i (for $1 \leq i \leq d$) involve also $\deg W$. Nonetheless, there is an isogeny of complexity n connecting a point of $W(\overline{\mathbb{Q}})$ to p with n bounded above by a polynomial in $h(p), [\mathbb{Q}(p) : \mathbb{Q}], \deg W$ and $[\mathbb{Q}(W) : \mathbb{Q}]$.

□

3.9 Avoiding problems in powers of modular curves

In this section we complete the proof of Theorem 3.6.7, that we state again for convenience.

Theorem. *Let a, b, c, r be positive integers and let $\mathcal{F} = \{X_1, X_2, \ldots\}$ be a set of subvarieties of $Y(1)^r$ of dimension $\leq a$, "algebro-geometric" degree $\leq b$ and field of definition of degree $\leq c$ over $\mathbb{Q}$. Let $Y \subseteq Y(1)^r$ be an irreducible subvariety such that for any positive integers i and n we have $Y \cap X_i^{(n)}$ is not Zariski dense in Y. Then, for d big enough, most points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $X_i(\overline{\mathbb{Q}})$ for any $X_i \in \mathcal{F}$.*

We now outline the scheme of the proof.

Let $p \in Y(\overline{\mathbb{Q}})$ be a point of degree d. Let $\mathcal{S}$ be a variety parameterising the subvarieties of $Y(1)^r$ of dimension a and degree b. We recall that a variety X of dimension a and degree b is defined over a number field K precisely when it is represented by a point of $\mathcal{S}(K)$.

1. We construct a definable set $S_{p,X}$ parameterising "virtual isogenies" between p and any point of $X(\mathbb{C})$; we require that integral points of $S_{p,X}$ correspond to isogenies between p and a point of $X(\mathbb{C})$. The sets $S_{p,X}$ form a definable family as p varies in $Y(\mathbb{C})$ and X (more precisely, its representing point) varies in $\mathcal{S}(\mathbb{C})$.

2. We try to classify the semialgebraic varieties contained in $S_{p,X}$: here the main obstruction in extending directly the proof in the basic case of $\mathcal{F}$ consisting of a family of lines appears, since in general some semialgebraic varieties can account for many points on X. Indeed, two points on $S_{p,X}$ are "connected" by a semialgebraic variety whenever they correspond to two points on X that lie on the same weakly special subvariety W. The converse, which

would be needed, is not quite true, for the following reason: given two weakly special subvarieties $W_1, W_2 \subseteq X$ with non-empty intersection, any point on $S_{p,X}$ corresponding to a point on W_1 can be "connected" to any any point on $S_{p,X}$ corresponding to a point on W_2 using a semialgebraic curve; we can "glue' (non smoothly) two semialgebraic curves exploiting a point on $S_{p,X}$ that corresponds to a point in the intersection $W_1 \cap W_2$.

3. In order to solve the issue, we can write $S_{p,X}$ as a finite disjoint union of definable subsets, keeping track of the special types of *maximal* weakly special subvarieties through a given point. By a result of Binyamini-Daw contained in [9], that we quote as our Theorem 3.9.3, any maximal weakly special subvariety of X has bounded "algebro-geometric" degree and by our explicit classification the maximal weakly special subvarieties of X then account for only finitely many special types. If Tp is a set of such special types, we denote with $S_{p,X,Tp}$ the (definable) subset of $S_{p,X}$ consisting of points that correspond to the points of X such that the special types of the maximal weakly special subvarieties of X through them are exactly the elements of Tp. , whose points correspond to the points of X which are contained in exactly n maximal weakly subvarieties of X. As $S_{p,X}$ is decomposed into finitely many subsets of the shape $S_{p,X,Tp}$, we can finally prove that whenever two points on $S_{p,X,Tp}$ are "connected" by a semialgebraic variety contained in $S_{p,X,Tp}$ then they correspond to points that lie on the same maximal weakly special subvarieties of X. All the estimates here are independent of p and X (as the bound on the degree of the maximal weakly special subvarieties given by the result of Binyamini-Daw depends on $\deg X$ only, which in turn is bounded by our hypotheses on $\mathcal{F}$).

4. We apply our form of isogeny estimates (Proposition 3.8.14) that roughly say that if p is isogenous to some point on a weakly special subvariety W, then we can bound the degree of an isogeny connecting p with some point on W with a polynomial in $d, h(p)$, the algebro-geometric degree of W and the degree of the field of definition of W. Since the algebro-geometric degree of W is uniformly (when $X \in \mathcal{F}$ varies) bounded whenever W is a maximal weakly special subvariety of X as a consequence of the result of Binyamini-Daw we have quoted above, we deduce that the degree of the field of definition of W is "big" in terms of the degrees of the connecting isogenies.

5. We can finally apply the Pila-Wilkie Theorem on each $S_{p,X,Tp}$, as the big degree of the field of definition of W produces many integral points that must be covered by many distinct semialgebraic varieties. This gives an upper bound on the degree of the field of definition of W and in turn we can bound the degree of the connecting isogenies.

6. Once we are done with bounding the degrees of the connecting isogenies, we bound the height of the point representing X in $\mathcal{S}$. This is obtained through an inductive argument. Essentially, if p is isogenous to $q \in X$ then either the Galois conjugates of q determine uniquely X in the family $\mathcal{F}$ and therefore X has bounded height, or there are many members of $\mathcal{F}$ that contain all the conjugates. However, we can always find a unique

element of $\mathcal{F}$ of *minimal* dimension containing all the conjugates of q, up to enlarging $\mathcal{F}$ suitably (e.g. if $\mathcal{F}$ contains all the surfaces of degree 2 then $\mathcal{F}$ will be extended so to include all the curves of degree 4, that can arise as intersections of the degree 2 surfaces).

7. We conclude the proof by a counting argument (where we use for the first time that $p \in Y$). As long as d is big enough, the number of points of degree $\leq d$ in $Y(\overline{\mathbb{Q}})$ with height bounded by $\log T$ will be comparable to T^{d^2+d}; the points on $Y(\overline{\mathbb{Q}})$ with height bounded by $\log T$ that can be isogenous to some $X \in \mathcal{F}$ parameterised by a point of $\mathcal{S}(\overline{\mathbb{Q}})$ of small height via isogenies of bounded degree turn out to be "sparse" and this proves Theorem 3.6.7.

Remark 3.9.1. *Before we start, a quick remark: we always assume that if $(p_1, p_2, \dots, p_r)$ and $(q_1, q_2, \dots, q_r)$ are isogenous then each p_i is isogenous to q_i for every $1 \leq i \leq r$. This might not always be the case, but we can easily reduce ourselves to this situation by enlarging the family $\mathcal{F}$ so that it includes all the X_i "with any permutation of coordinates".*

Let us now perform the steps outlined above.

1. Let $p = (p_1, p_2, \dots, p_r)$. We define $S_{p,X}$ as

$$S_{p,X} = \{(a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, \dots, a_r, b_r, c_r, d_r) \in \mathbb{R}^{4r} \mid \exists q \in X(\mathbb{C})$$
$$\left(\frac{a_1 j^{-1}(p_1) + b_1}{c_1 j^{-1}(p_1) + d_1}, \frac{a_2 j^{-1}(p_2) + b_2}{c_2 j^{-1}(p_2) + d_2}, \dots, \frac{a_r j^{-1}(p_r) + b_r}{c_r j^{-1}(p_r) + d_r}\right) = j^{-1}(q)\}$$

   where, as usual, j is restricted to the fundamental domain $D \subseteq \mathbb{H}$.

2. We have outlined already, in the previous sections, some semialgebraic varieties contained in $S_{p,X}$, namely those that correspond to a fixed q.
   In general the situation is more complicated and indeed some semialgebraic varieties contained in $S_{p,X}$ can account for more than one q. This happens precisely when X contains positive dimensional weakly special subvarieties; we quote the hyperbolic version of Ax-Lindemann-Weierstrass proven by Klingler, Yafaev and Ullmo [34] adapted to this context.
   We define $j(M)$, for any $M \in S_{p,X}$, by

$$j(M) = j(M \cdot j^{-1}(p)).$$

   For example, $j(S_{p,X})$ consists exactly of the admissible $j^{-1}(q)$ arising in the definition of $S_{p,X}$ above.

   Theorem 3.9.2 (Hyperbolic Ax-Lindemann-Weierstrass)**.** *Let V be a semialgebraic subvariety contained in $S_{p,X}$. Any irreducible component of the Zariski closure of $j(V) \subseteq X$ is a weakly special subvariety and if $W \subseteq X$ is a weakly special subvariety then $j^{-1}(W)$ is a semialgebraic subset of $S_{p,X}$.*

In this sense two points on $S_{p,X}$ are "connected" by a connected semialgebraic variety whenever they correspond to two points on the same weakly special subvariety. We would really like an inverse statement, in order to perform the following heuristic: suppose that $q \in W$, where $W \subseteq X$ is a positive dimensional weakly special subvariety; if the degree of the field of definition of W is big, we would find many Galois conjugates of W and we could hope that for some of these, say W', and a Galois conjugate of q, say q', on W' there was no semialgebraic variety "connecting" the points in $S_{p,X}$ representing q and q'. This is not the case (and in this context the Pila-Wilkie Theorem 2.2.5 would just not give the required information - we need no connected semialgebraic variety to account for both q and q'), as we can observe with the following example. Suppose that two positive dimensional weakly special subvarieties $W_1, W_2 \subseteq X$ have a non-empty intersection and let $w \in W_1 \cap W_2$ be a point. The point in $S_{p,X}$ that represents q_1 (resp. q_2) can be "connected" to that which represents w (by Theorem 3.9.2), as $q_1, w \in W_1$ (resp. $q_2, w \in W_2$). Thus, we can find two semialgebraic curves, one accounting for q_1 and w, the other accounting for q_2 and w, and "glue" them (most likely in a non-smooth way) at the intersection.
The key observation here is that indeed w will be contained in "more" weakly special subvarieties than q_1 and q_2 (at least if they are "randomly chosen" points of W_1 and W_2). We prove the correct statement in the next step.

3. We now decompose $S_{p,X}$ into suitable definable subsets that take into account the special types of the maximal weakly special subvarieties through a point. We first need a result by Binyamini-Daw [9], that we rephrase here (in a weaker form) for our purposes:

   Theorem 3.9.3. *There is a polynomial $P(t)$, depending on r only, such that for every subvariety $X \subseteq Y(1)^r$ and every maximal weakly special subvariety $W \subseteq X$ we have*

   $$\deg W < P(\deg X).$$

   Remark 3.9.4. *The result by Binyamini-Daw is indeed much stronger: a similar bound holds for all the weakly optimal subvarieties of X.*

   As an immediate application, any point of X is contained in a bounded number of maximal weakly special subvarieties (of X): the admissible degrees of these depend on $\deg X$ only and, by our explicit description, there is an explicit upper bound on the amount of admissible special types of these weakly special subvarieties.
   Moreover, the set of special types of maximal weakly special subvarieties of X is contained in a finite set TP which is independent of p and X. Indeed, the degree of $X \in \mathcal{F}$ is uniformly bounded by our hypotheses on $\mathcal{F}$ and in turn the degrees of the maximal special subvarieties are uniformly bounded as well (and they account for only finitely many special types).

Let then Tp be a set whose elements are special types chosen among the elements of TP, i.e. among all the possible special types of maximal weakly special subvarieties of X as X varies in $\mathcal{F}$. We define

$$S_{p,X,Tp} = \{z \in S_{p,X} \mid \ z \text{ corresponds to a point on } X$$

$$\text{contained in exactly } |Tp| \text{ maximal weakly special subvarieties of } X,$$

$$\text{each with special type an element of } Tp\}.$$

The set formalism for Tp (rather than asking for a multi-set) is sufficient: two distinct weakly special subvarieties of X with the same special type are disjoint and we observe that, as Tp varies among the subsets of TP, the disjoint finite union of the sets $S_{p,X,Tp}$ is $S_{p,X}$.

We observe that $S_{p,X,Tp}$ is definable (uniformly in p, X): as the maximal weakly special subvarieties of X have bounded degree, they are contained in a definable family of varieties (this can be seen very explicitly by our description - the weakly special subvarieties of $Y(1)^r$ of a given special type form a definable family) and (uniform in p, X) definability of $S_{p,X,Tp}$ follows. As an example, the subset of $S_{p,X}$ consisting of points that correspond to points on X contained in a maximal weakly special subvariety of a given type are described by the formula

$z \in S_{p,X}$ such that there is a weakly special subvariety W of such given type, which:

(a) is a subset of in X;

(b) contains the point of X corresponding to z;

(c) is maximal among all the weakly special subvarieties contained in X (i.e. one just checks that for no W' chosen among the weakly special subvarieties of $Y(1)^r$ of the admissible special types we have have $W \subseteq W' \subseteq X$ unless $W = W'$).

The set $S_{p,X,Tp}$ is an appropriate conjunction of formulae of this shape (and their negation).

We fix a (finite) subset Tp of TP and we can finally prove the following.

Let $q, q' \in X$ such that there is a connected semialgebraic subvariety V of $S_{p,X,Tp}$ accounting for both points (i.e. V contains two points corresponding to q and q'). Then, for any maximal weakly special subvariety $W \subseteq X$, we have that $q \in W$ if and only if $q' \in W$.

Remark 3.9.5. *This is not obvious in principle, as decomposing $S_{p,X}$ according to the special types of maximal weakly special subvarieties of X is not the same as decomposing it according to the maximal weakly special subvarieties of X themselves. As an example, let us consider the product of a non-weakly special curve $E \subseteq Y(1)^m$ and a special subvariety*

$Z \subseteq Y(1)^n$: the product $E \times Z \subseteq Y(1)^{m+n}$ consists of points that are contained each in a single maximal weakly special subvariety (i.e. a copy of Z), so the decomposition above would give us no information in that context. However, the statement above would still be true. This example also shows that in general there can be infinitely many maximal weakly special subvarieties of X (in contrast to the finiteness of the set of their special types).

Let $\alpha : [0,1] \to S_{p,X,Tp}$ be a semialgebraic curve such that $\alpha(0)$ corresponds to $q \in W$, with $W \subseteq X$ maximal weakly special.
Let $C_1 \times C_2 \times \ldots \times C_d$ be the special type of W. Since $\alpha(t) \in S_{p,X,Tp}$ for every $t \in [0,1]$, there are maximal weakly special subvarieties $W_t \subseteq X$ of the shape

$$C_1 \times C_2 \times \ldots \times C_d \times P_t \subseteq Y(1)^{r_1+r_2+\ldots+r_d} \times Y(1)^{r'}$$

such that $\alpha(t) \in W_t$ for every $t \in [0,1]$. Let us now recall our definition of $S_{p,X}$ and notice that the (possibly non-algebraic) curve defined as the image of

$$\beta : [0,1] \to Y(1)^{r'}$$

$$\beta(t) = P_t$$

is the image via j of a semialgebraic curve contained in $D^{r'}$. This prompts us to use some Ax-Schanuel type result and indeed we can use again the Hyperbolic Ax-Lindemann-Weierstrass Theorem 3.9.2, which in this case implies that the Zariski closure of $\beta([0,1]) \subseteq Y(1)^{r'}$ is a weakly special subvariety W'. We claim that

$$C_1 \times C_2 \times \ldots \times C_d \times W' \subseteq X.$$

Indeed, the set

$$\{P \in Y(1)^{r'} \mid C_1 \times C_2 \times \ldots \times C_d \times P \subseteq X\}$$

is Zariski closed, as it is the complement of the projection (which is an open map) on the last r' coordinates of

$$C_1 \times C_2 \times \ldots \times C_d \times Y(1)^{r'} \setminus (C_1 \times C_2 \times \ldots \times C_d \times Y(1)^{r'} \cap X) \subseteq Y(1)^r.$$

As

$$C_1 \times C_2 \times \ldots \times C_d \times \beta([0,1]) \subseteq X$$

the claim follows. Therefore

$$W \subseteq C_1 \times C_2 \times \ldots \times C_d \times W' \subseteq X$$

and, by maximality of W, we have

$$W = C_1 \times C_2 \times \ldots \times C_d \times W'.$$

Observe that W is just

$$C_1 \times C_2 \times \ldots \times C_d \times \beta(0)$$

i.e. $\beta : [0,1] \to Y(1)^{r'}$ is a constant curve and thus the image of $\alpha : [0,1] \to S_{p,X,Tp}$ accounts only for points of X contained in W. This proves the statement above for connected semialgebraic curves (and indeed any connected semialgebraic variety) in $S_{p,X,Tp}$.

4. We now apply our form of isogeny estimates taking into account positive dimensional weakly special subvarieties, namely Proposition 3.8.14.
In our context, as we have deduced above as a consequence of the Binyamini-Daw results, the "algebro-geometric" degree of any maximal weakly special subvariety with respect to any member of $\mathcal{F}$ is bounded (since the degrees of its members are bounded). We therefore obtain, assuming that p is isogenous to some point of $X(\overline{\mathbb{Q}})$ contained in a maximal weakly special subvariety $W \subseteq X$:

$$p \text{ is isogenous to some point of } W(\overline{\mathbb{Q}}) \text{ via an isogeny of complexity at most}$$

$$P(h(p), [\mathbb{Q}(p) : \mathbb{Q}], [\mathbb{Q}(W,p) : \mathbb{Q}(p)])$$

$$\text{for a polynomial } P(t_1,t_2,t_3) \in \mathbb{Q}[t_1,t_2,t_3] \text{ depending on } r \text{ and } \mathcal{F} \text{ only.}$$

We have replaced $[\mathbb{Q}(W) : \mathbb{Q}]$ with $[\mathbb{Q}(W,p) : \mathbb{Q}(p)]$ for convenience; this is legitimate as

$$[\mathbb{Q}(W,p) : \mathbb{Q}(p)] \leq [\mathbb{Q}(W) : \mathbb{Q}]$$

and

$$[\mathbb{Q}(W) : \mathbb{Q}] \leq [\mathbb{Q}(W,p) : \mathbb{Q}(p)][\mathbb{Q}(p) : \mathbb{Q}].$$

5. Suppose that p is isogenous to some point of $X(\overline{\mathbb{Q}})$ lying on W for a maximal weakly special subvariety $W \subseteq X$ (which can possibly just be the point itself). We argued in 4. that there is some $q \in W(\overline{\mathbb{Q}})$ such that p and q are related by an isogeny of complexity n satisfying

$$n < P(h(p), [\mathbb{Q}(p) : \mathbb{Q}], [\mathbb{Q}(W,p) : \mathbb{Q}(p)])$$

for a polynomial $P(t_1,t_2,t_3) \in \mathbb{Q}[t_1,t_2,t_3]$ which depends on r and $\mathcal{F}$ only. By our definition of complexity of an isogeny and the construction of $S_{p,X}$, this means that we have produced an integral point on $S_{p,X}$ whose exponential height is bounded above by n, a number which is bounded polynomially in terms of $h(p)$, $[\mathbb{Q}(p) : \mathbb{Q}]$ and $[\mathbb{Q}(W,p) : \mathbb{Q}(p)]$. Let us consider the $[\mathbb{Q}(W,p) : \mathbb{Q}(p)]$ Galois conjugates (over $\mathbb{Q}(p)$) of W; we recall that, by our explicit description, they are all disjoint. Each of these Galois conjugations gives rise to a point $q' \in W'(\overline{\mathbb{Q}})$ for some conjugate W' of W and moreover q' is related to p by an isogeny of complexity n (since the modular polynomials have integral coefficients).

If X is defined over $\mathbb{Q}(p)$, all these $[\mathbb{Q}(W,p):\mathbb{Q}(p)]$ many conjugates of q give rise to integral points on $S_{p,X}$; in general, at least

$$[\mathbb{Q}(W,p,X):\mathbb{Q}(p,X)] \geq [\mathbb{Q}(W,p):\mathbb{Q}(p)]/[\mathbb{Q}(X):\mathbb{Q}]$$

many of the Galois conjugations considered above fix X and, as the degree of the field of definition of X is bounded above by our hypotheses on $\mathcal{F}$, there is an explicit constant $e > 0$ such that we indeed produce at least $e \cdot [\mathbb{Q}(W,p):\mathbb{Q}(p)]$ integral points on $S_{p,X}$, each accounting for different maximal special subvarieties.
Decomposing $S_{p,X}$ as a finite disjoint union of sets of the shape $S_{p,X,Tp}$ as above, we see that the $e \cdot [\mathbb{Q}(W,p):\mathbb{Q}(p)]$ many integral points we have just produced cannot be "connected" by semialgebraic sets as they belong to disjoint maximal weakly special subvarieties and finally the Pila-Wilkie Theorem 2.2.5 reads:

for every $\varepsilon > 0$, there is a positive constant c such that any collection of integral points of exponential height bounded above by n accounting for different maximal weakly special subvarieties on $S_{p,X}$ is of size at most $c \cdot n^{\varepsilon}$.

This implies that

$$e \cdot [\mathbb{Q}(W,p):\mathbb{Q}] < c \cdot P(h(p), [\mathbb{Q}(p):\mathbb{Q}], [\mathbb{Q}(W,p):\mathbb{Q}])^{\varepsilon}$$

and choosing ε^{-1} bigger than the degree of $P(t_1,t_2,t_3)$ in t_3, we finally bound the complexity of the connecting isogenies in terms of p only:

$$e \cdot [\mathbb{Q}(W,p):\mathbb{Q}] < P(h(p), [\mathbb{Q}(p):\mathbb{Q}])$$

$$n < P(h(p), [\mathbb{Q}(p):\mathbb{Q}])$$

where $P(t_1,t_2) \in \mathbb{Q}[t_1,t_2]$ is an appropriate polynomial depending on r and $\mathcal{F}$ only.

6. We have that if $p \in Y(\overline{\mathbb{Q}})$ is isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$, then there is some $q' \in X(\overline{\mathbb{Q}})$ such that p and q' are related by an isogeny of complexity n, with

$$n < P(h(p), [\mathbb{Q}(p):\mathbb{Q}])$$

where $P(t_1,t_2) \in \mathbb{Q}[t_1,t_2]$ is an appropriate polynomial depending on r and $\mathcal{F}$ only.

We wish to prove that the isogenous images $X^{(n)}$ of X, where $X^{(n)}$ denotes the subvariety of $Y(1)^r$ consisting of points that are related to a point of X by an isogeny of complexity n, are "far" from covering Y.
In order to do this, we need first to bound the height of X (namely, the height of the point that corresponds to X in the variety $\mathcal{S}$ that parameterises subvarieties of $Y(1)^r$ of a given degree) in terms of $h(p)$ and $[\mathbb{Q}(p):\mathbb{Q}]$.
The idea is the following: suppose that $q \in X(\overline{\mathbb{Q}})$ has a field of definition of big degree.

Then we can produce many conjugates of q always lying on X and hopefully bound the height of the point corresponding to X in $\mathcal{S}$: for instance, as we have seen already, if X is a line then it is uniquely determined by two points and we can express the height of the coefficient of a defining equation in terms of the height of q. Notice that the conjugates of q will need not be required to be isogenous to p anymore now. It is fundamental that X is uniquely determined in a suitable family of varieties and we can apply Lemma 3.7.8.

Let b be an upper bound for the "algebro-geometric" degree of any element X_i of $\mathcal{F}$ and we recall that $\mathcal{S}$ is a variety parameterising all the subvarieties of $Y(1)^r$ of degree at most b, so that each X_i is represented by some element of $\mathcal{S}$. In the subsequent lines, we treat $\mathcal{S}$ (which is a parameter space) also just as the set of varieties its points represent and we refer to an element of $\mathcal{S}$ (a point) as the subvariety of $Y(1)^r$ it represents.
We now define $\mathcal{V}$ as

$$V \in \mathcal{V}$$

if and only if

$$V \text{ is a finite intersection of elements represented by } \mathcal{S}.$$

Remark 3.9.6. *An arbitrary intersection of varieties is just a finite intersection by Hilbert's Basis Theorem, so $\mathcal{V}$ is just the set of intersections of varieties parameterised by $\mathcal{S}$*

Remark 3.9.7. *We might also assume that $Y(1)^r \in \mathcal{V}$, as the "empty intersection".*

We prove that $\mathcal{V}$ is indeed a variety, i.e. the finite intersections of elements of $\mathcal{S}$ form a "continuous family". We first introduce a notion to keep track of the degrees of reducible varieties.
Let $V \subseteq Y(1)^r$ be a variety. We define $\sigma(V)$, the *signature* of V, as the sequence $(a_r, a_{r-1}, \ldots, a_0)$ where a_i is the sum of the degrees of the irreducible components of V of dimension i.
We use the lexicographic order on these signatures, in this sense:

$$(a_r, a_{r-1}, \ldots, a_0) < (a'_r, a'_{r-1}, \ldots, a'_0)$$

if and only if for some natural number $0 \leq k \leq r$ we have

$$a_i = a'_i \text{ for every } i > k$$

and

$$a_k < a'_k.$$

A nice property of the signature is that whenever $V' \subseteq V$ then

$$\sigma(V') \leq \sigma(V).$$

This is a (total) well order on the signatures, with strict equality if V' (which is closed) is strictly contained in V.

Lemma 3.9.8. *Let $\mathcal{T}'$ and $\mathcal{T}$ be two families of varieties (in the sense that they are subvarieties of some common Chow variety). We prove that there is a positive integer N, depending on $\mathcal{T}'$ and $\mathcal{T}$, such that any intersection*

$$T' \cap T_1 \cap T_2 \cap \ldots \cap T_k$$

for some $T' \in \mathcal{T}'$ and some $T_1, T_2, \ldots, T_k \in \mathcal{T}$ is equal to an intersection

$$T' \cap T_{i_1} \cap T_2 \cap \ldots \cap T_{i_n}$$

for some appropriate $n \leq N$ and positive integers $i_1 \leq i_2 \leq \ldots \leq i_n \leq k$.

This means that any arbitrary finite intersection is actually an intersection of a uniformly bounded number of members.

Proof. Since $\mathcal{T}'$ is a family of varieties of bounded degree, only finitely many signatures of its elements are allowed. We prove the statement above by induction on the maximum of such signatures, in the sense of the lexicographic (well) order we have put on signatures. Our claim is obvious if $\mathcal{T}'$ parameterises only a bounded a number of points. Let $T' \in \mathcal{T}'$ be an element of maximum signature and we suppose that our claim holds for any family whose maximum signature of elements is strictly smaller than $\sigma(T')$. We consider the intersection

$$T' \cap T_1 \cap T_2 \cap \ldots \cap T_k$$

and we first note that either $T' \subseteq T_i$ for every $1 \leq i \leq k$, or some intersection, say without loss of generality setting $i = 1$,

$$T' \cap T_1$$

is a proper subvariety of T', so that $\sigma(T' \cap T_1) < \sigma(T')$. We observe that the subset of $\mathcal{T}'$ consisting of the subvarieties with signature strictly smaller than $\sigma(T')$ is indeed a subvariety of $\mathcal{T}'$, since the statement

$$t' \in \mathcal{T}' \text{ parameterises a subvariety of signature } \sigma(T)$$

is definable over $\mathbb{C}$ (in the language of the theory of rings) in $\mathcal{T}'$.

Then we can apply our inductive hypothesis to

$$(T' \cap T_1) \cap T_2 \cap T_3 \cap \ldots \cap T_k$$

and obtain an intersection of a bounded number of terms. Notice that such bound is uniform as it depends only on a fixed subfamily of $\mathcal{T}'$. □

As an immediate consequence of this Lemma, we see that $\mathcal{V}$ consists of elements that satisfy a common upper bound on their degree, as

$$\deg_c(T_1 \cap T_2 \cap \ldots \cap T_k) \leq \deg_c T_1 \cdot \deg_c T_2 \cdot \ldots \cdot \deg_c T_k$$

for any $T_1, T_2, \ldots, T_k \in Y(1)^r$. Thus, $\mathcal{V}$ can be seen as a subset of an appropriate Chow variety and the property (depending on k)

$$V \in \mathcal{V} \text{ is an intersection of at most } k \text{ elements of } \mathcal{S}$$

is definable over $\mathbb{C}$ (in the language of the theory of rings). Hence, since the number of terms in the intersection is uniformly bounded by the Lemma above, setting $\mathcal{T}' = \mathcal{T} = \mathcal{S}$, we obtain that $\mathcal{V}$ can indeed be chosen as a variety.

We observe that $\mathcal{V}$, when embedded in the appropriate Chow variety, is even defined over $\mathbb{Q}$, since $\mathcal{S}$ is. We are just asserting that the set of arbitrary intersections of varieties of degree at most b is stable by Galois conjugation.

Let us now explain what comes next. As $p \in Y(\overline{\mathbb{Q}})$, isogenous to some $q \in X(\overline{\mathbb{Q}})$, with $X \in \mathcal{F}$, we set K as the field of definition of X. All the Galois conjugates of q over K, say $q_1, q_2, \ldots, q_k$, lie on $X(\mathbb{C})$ and we could hope that $q_1, q_2, \ldots, q_k$ determine uniquely X, in the sense that X is the minimal subvariety represented in $\mathcal{S}$ that contains such points. In this case, we could at least in principle - notice that there is an issue that we describe with below, concerning the magnitude of k - bound the height of the point representing X in $\mathcal{S}$ using Lemma 3.7.8.
This might not always be the case, but there is indeed a minimal variety represented by $\mathcal{V}$ that contains all these points: we just take the intersection of all the elements of $\mathcal{V}$ that contain $q_1, q_2, \ldots, q_k$. This variety, say V, is indeed defined over K, since any Galois conjugation fixing K permutes the point $q_1, q_2, \ldots, q_k$ and if V and V' were distinct but related by Galois conjugation, then $V \cap V'$ would be a proper subvariety of V, contradicting its minimality.
Thus, we can indeed take a unique variety containing the points $q_1, q_2, \ldots, q_k$. We cannot apply Lemma 3.7.8 uniformly yet, since the number k of conjugates of q over K depends on the degree of the field of definition of q. Although in 5. above we concluded with an upper bound on the degree of the connecting isogenies and we can obtain bounds on the degree of the field of definition of q, in order to use Lemma 3.7.8 we need to bound absolutely k. This amounts to the following statement.

Lemma 3.9.9. *Let $\mathcal{T}$ be a family of subvarieties closed by intersection. There is a positive integer M, depending on $\mathcal{T}$ only, such that, for any finite sequence of points $q_1, q_2, \ldots, q_k$, if T is the minimal element of $\mathcal{T}$ such points, there are positive integers $1 \leq i_1 \leq i_2 \leq \ldots \leq i_n \leq k$ such that T is the minimal element of $\mathcal{T}$ containing the points $q_{i_1}, q_{i_2}, \ldots, q_{i_M}$.*

Remark 3.9.10. *In the case of a family of lines this is immediate: a line is determined by two points and, if we only have one points, we just choose that very point.*

Proof. By lexicographic induction on the maximum signature among the signatures of the elements of $\mathcal{T}$. If the elements of $\mathcal{T}$ are just a uniformly bounded amount of points, the result is immediate.
Suppose that $T \in \mathcal{T}$ has maximum signature. We will consider the subfamily of $\mathcal{T}$ obtained by excluding the elements of signature $\sigma(T)$ that, as we argued above, in indeed parameterised by a subvariety of $\mathcal{T}$, which is closed by intersection.
We argue that, up to permuting the indices of $q_1, q_2, \ldots, q_k$, there is some j such that the minimal element of $\mathcal{T}$ containing $q_1, q_2, \ldots, q_j$ is a proper subvariety of T, while, for every $j+1 \leq i \leq k$, the minimal element of $\mathcal{T}$ containing the points $q_1, q_2, \ldots, q_j, q_i$ is T. This can be attained choosing a maximal value of j and, if there is no such value, even a single point would determine T as the minimal element of $\mathcal{T}$ containing it.
Let T' be the minimal variety of $\mathcal{T}$ containing $q_1, q_2, \ldots, q_j$. Since T' is a proper subvariety of T, we have that T' is an element of the subfamily $\mathcal{T}'$ of $\mathcal{T}$ defined as $\mathcal{T}$ excluding the elements of signature $\sigma(T)$. We can therefore apply our inductive hypothesis on such family and we obtain that for some m (depending on $\mathcal{T}$ only), the variety T' is the minimal element of $\mathcal{T}'$ containing the points $q_{i_1}, q_{i_2}, \ldots, q_{i_m}$ for appropriate $1 \leq i_1 \leq i_2 \leq \ldots \leq i_m \leq j$. We notice that T' is also the minimal element $\mathcal{T}$ containing $q_{i_1}, q_{i_2}, \ldots, q_{i_m}$, since $T' \cap V \in \mathcal{T}'$ for any $V \in \mathcal{T}$.
We argue that T is the minimal element of $\mathcal{T}$ containing $q_{i_1}, q_{i_2}, \ldots, q_{i_m}, q_{j+1}$. Indeed, any element of $\mathcal{T}$ containing $q_{i_1}, q_{i_2}, \ldots, q_{i_m}$ must contain T' by minimality and thus it must contain $q_1, q_2, \ldots, q_j$. We choose j so that the minimal element of $\mathcal{T}$ containing $q_1, q_2, \ldots, q_j, q_{j+1}$ and this proves our claim. The proof follows setting $M = m + 1$. □

We are now ready to complete the argument: if $p \in Y(\overline{\mathbb{Q}})$ is isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$ defined over the number field K, there is some uniquely determined minimal variety (which is a subvariety of X) parameterised by a point of $\mathcal{V}(K)$ containing all the conjugates $q_1, q_2, \ldots, q_k$. Moreover, such variety is determined by only M points among the $q_1, q_2, \ldots, q_k$ and we can finally apply Lemma 3.7.8. We obtain the following, noting that all the heights $h(q_1), h(q_2), \ldots, h(q_k)$ are equal to $h(q)$ by Galois invariance of the height.

Let $p \in Y(\overline{\mathbb{Q}})$ be isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$ defined over the number field K. There is a point $z \in \mathcal{V}(K)$ such that q is contained in the variety represented by z and

$$h(z) < c(1 + h(q))$$

for an appropriate constant $c > 0$ depending on $\mathcal{F}$ only.

We remark that such variety represented by z will indeed be a subvariety of X.

7. We are now ready to prove that "most" points of $Y(\overline{\mathbb{Q}})$ are not isogenous to any point of any $X_i(\overline{\mathbb{Q}})$. For any positive integers $T \geq 10$ and $d \geq 2$ let

$$Y_{T,d} = \{p \in Y(\overline{\mathbb{Q}}) \mid h(p) < \log T \text{ and } [\mathbb{Q}(p) : \mathbb{Q}] \leq d\}.$$

We think of d as a big fixed positive integer and we claim that, as long as T is big enough, $Y_{T,d}$ is "far" from being covered by the varieties parameterised by $\mathcal{V}$ we obtained in 6. and their isogenous images of an appropriately small degree.
We make this precise. Suppose that $p \in Y_{T,d}$ is isogenous to some $q \in X$ for some $X \in \mathcal{F}$. By the conclusion of 5., there is some $q' \in X(\overline{\mathbb{Q}})$ such that p is connected to q' by an isogeny of complexity n and the inequality

$$n < P(\log T, d)$$

holds for an appropriate polynomial $P(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ depending on $\mathcal{F}$ only.
On the other hand, by the conclusion of 6. there is a subvariety $Z \subseteq X$ such that Z is parameterised by a point z of $\mathcal{V}(K)$, where K is the field of definition of X, such that:

$$\begin{gathered} q' \in Z(\overline{\mathbb{Q}}) \\ h(z) < c_1(1 + h(q')) \end{gathered}$$

for an appropriate constant $c_1 > 0$ depending on $\mathcal{F}$ only.
The height of q' can be bounded by using Pazuki's Theorem 3.5.4 on each individual coordinate, as q' and p are connected by an isogeny of complexity n, and we obtain

$$h(q') < c_2(1 + h(p) + \log n) < c_3(1 + \log T + \log\log T + \log d)$$

for appropriate constants c_2, c_3 depending on $\mathcal{F}$ only (we are using the conclusion of 5. recalled right above). Thus, for some constant c_4 depending on $\mathcal{F}$ only we have

$$h(z) < c_4(1 + \log T + \log d)$$

and we have that the degree of z over $\mathbb{Q}$ is absolutely bounded as $z \in \mathcal{V}(K)$, with K the field of definition of X (which has degree over bounded by our hypotheses on $\mathcal{F}$). Namely, the bound from the statement of Theorem 3.6.7 is

$$[\mathbb{Q}(z) : \mathbb{Q}] \leq c.$$

Let us now count the points of $\mathcal{V}(\overline{\mathbb{Q}})$ with bounded height and bounded degree and we define as above

$$V_{U,e} = \{z \in V(\overline{\mathbb{Q}}) \mid h(z) < \log U \text{ and } [\mathbb{Q}(z) : \mathbb{Q}] \leq e\}.$$

We wish to estimate from above the size of $V_{U,c}$ for

$$U = \exp\left(c_4(1 + \log T + \log d)\right).$$

We just need to know bounds on projective spaces: since $\mathcal{V}$ is a subset of a Chow variety, then it can be embedded in a suitable projective space (with an embedding defined over $\mathbb{Q}$). Notice that, even if $\mathcal{V}$ was not quasi-projective, the following argument would hold by splitting $\mathcal{V}$ into finitely many affine varieties.
By the estimates of Masser and Vaaler in [40] that we discussed above, we obtain

$$|V_{U,c}| < c_5 \cdot U^{c^2 \cdot c_6}$$

for appropriate constants c_5 and c_6 depending on $\mathcal{V}$ only (and hence on $\mathcal{F}$ only). The constant c_6 takes into account the dimension of the projective space we are embedding $\mathcal{V}$ into.
Thus, the $z \in \mathcal{V}(\overline{\mathbb{Q}})$ satisfying the two bounds

$$[\mathbb{Q}(z) : \mathbb{Q}] \leq c \text{ and } h(z) < c_4(1 + \log T + \log d)$$

are at most

$$s(T, d)$$

many, for an appropriate polynomial $s(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ depending on $\mathcal{F}$ only. Recall that c is an absolute constant implied by the statement of Theorem 3.6.7.

Each of the varieties parameterised by a point of $\mathcal{V}$ has an "algebro-geometric" degree bounded above by b', a quantity which depends on $\mathcal{F}$ only.
We can finally summarise the observations above: if any $p \in Y_{T,d}$ is isogenous to some point of $X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$, then it is isogenous, via an isogeny of complexity n, to some $q' \in X(\overline{\mathbb{Q}})$ contained into some $Z \subseteq X$ parameterised by a point of $z \in \mathcal{V}(\overline{\mathbb{Q}})$ that is chosen in a set (depending on $\mathcal{F}, T, d$ only - in particular, independent of X) of size at most $s(T, d)$.
The union of all these admissible varieties Z is a (possibly reducible) variety Z' of degree at most

$$b' \cdot s(T, d)$$

so that such degree grows at most polynomially in T and d.
We notice that no isogenous image $Z'^{(n)}$ will cover Y, where, given a subset $S \subseteq Y(1)^r$, we denote with $S^{(n)}$

$$S^{(n)} = \{s \in Y(1)^r \mid s \text{ is isogenous to some element of } S$$

$$\text{via an isogeny of complexity } n\}$$

since $Y \not\subseteq X^{(n)}$ for any $X \in \mathcal{F}$ and any positive integer n and Z' is a finite union of subvarieties of some $X_i \in \mathcal{F}$.
We recall the bound on n, which is

$$n < P(\log T, d)$$

for an appropriate polynomial $P(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ depending on $\mathcal{F}$ only. As above, the degree of $Z'^{(n)}$ can be deduced from Merten's Theorem (Remark 3.4.8) and the arguments by Masser-Zannier in [42]. We obtain

$$\deg Z'^{(n)} \leq c_7 \cdot \deg Z \cdot (n \log n)^{2r}$$

for some constant c_7 depending on r only and then

$$\deg Z'^{(n)} \leq c_7 \cdot b' \cdot s(T, d) \cdot (P(\log T, d) \cdot \log P(\log T, d))^{2r}.$$

Finally, summing over the values of $n < P(\log T, d)$ and intersecting with Y, we obtain that any $p \in Y_{T,d}$ which is isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$ is contained in the proper subvariety

$$Y' = Y \cap \bigcup_{n=1}^{P(\log T, d)} Z'^{(n)} \subseteq Y$$

of degree bounded above by $q(T, d)$, where $q(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ is an appropriate polynomial depending on $\mathcal{F}$ only.
Notice that Y' might indeed depend on T and d, but its degree will grow at most like a fixed power of T and d.

This is sufficient to conclude provided d is big enough and we now show that Y' is "far" from covering $Y_{T,d}$. We reduce us to the case of $Y \cong \mathbb{A}^m$ (with $m = \dim Y$) using the following argument.
Let $U \subseteq Y$ be a dense open subset that admits a dominant quasi-finite morphism (i.e. with finite fibers) φ defined over $\overline{\mathbb{Q}}$ to some affine space $\mathbb{A}^m$; we can just restrict to an open affine subset of Y and apply Noether's normalisation Lemma.
The degree of the image of $U \cap Y'$ is again bounded by a fixed polynomial of T and d, by the usual properties of the classical degree: namely, the graph $\Gamma(\varphi)$ of the morphism φ described above is a subvariety of $U \times \mathbb{A}^m$ and the image $\varphi(U \cap Y')$ is just the projection on $\mathbb{A}^m$ of the intersection of $\Gamma(\varphi)$ with $(U \cap Y' \times \mathbb{A}^m)$ in $U \times \mathbb{A}^m$. We repeatedly use that the degree of an intersection of two varieties is bounded above by the product of the degrees of the two varieties and that the degree of a projection is bounded above by the degree of the former variety.
Let k be a number field over which Y, U and φ are defined. Since φ has finite fibers over every point of $\mathbb{A}^m$, there is an explicit constant $c_8 > 0$ (depending on Y, U, φ, k only) such that

$$c_8^{-1} \cdot [\mathbb{Q}(q) : \mathbb{Q}] < [\mathbb{Q}(p) : \mathbb{Q}] < c_8 \cdot [\mathbb{Q}(q) : \mathbb{Q}]$$

for every $p \in U(\overline{\mathbb{Q}})$ such that $q = \varphi(p)$. Analogously, possibly shrinking U, we can apply Theorem 2.7.2 and obtain that for some constant $c_9 > 0$ (depending on Y, U, φ, k only) then

$$c_9^{-1} \cdot (1 + h(q)) < (1 + h(p)) < c_9 \cdot (1 + h(q)).$$

These considerations essentially add up to observing that any relevant quantity ("algebro-geometric" degrees, degrees of the fields of definition and heights) only varies by a fixed bounded amount when we deal with $\mathbb{A}^m$ rather than Y.

Let us then prove that, for any fixed $c_{10} > 0$, most points of $\mathbb{A}^m$ of height bounded above by $c_{10} \cdot \log T$ and degree bounded above by $c_{10} \cdot d$ are not contained in the image H of $Y' \cap U$ (when d is big enough). We point out that the additional constant c_{10} has the sole purpose of allowing us to "track back" points to Y.
Let us consider "vertical" hyperplanes of $\mathbb{A}^m$, by which we mean hyperplanes which are determined by the value of the first coordinate. These hyperplanes are parameterised by a variety G isomorphic to $\mathbb{A}^1$ (that we identify with the first coordinate of $\mathbb{A}^m$) and the subset of G parameterising hyperplanes which are not contained in the closure of H is open: this must be the case since the subset of G parameterising hyperplanes contained in the closure of H is a definable (over $\mathbb{C}$, in the language of rings) subset of G and if such definable set was not contained in a Zariski closed subset of G then H would be dense in $\mathbb{A}^m$ (but Y' is not dense in Y).
Thus, only finitely many vertical hyperplanes are contained in H. Each of these is an irreducible subvariety of H (since H is a proper subvariety of $\mathbb{A}^m$) and therefore there are at most $\deg H$ many of such. Using again the estimates of Masser and Vaaler, the size of

$$\{\alpha \in \mathbb{A}^1(\overline{\mathbb{Q}}) \mid h(\alpha) < \log T \text{ and } [\mathbb{Q}(\alpha) : \mathbb{Q}] \leq d\}$$

is comparable to T^{d^2+d}. Thus, if d is bigger than the degree in t of the polynomial, not depending on T and d, that gives an upper bound on $\deg H$ for any choice of H, T, d, then for T big enough "most" points of G of degree bounded by d and height bounded by $\log T$ will correspond to an hyperplane which is not contained in H.
The intersection of a vertical hyperplane with H produces some $H' \subseteq \mathbb{A}^{m-1}$ of the same degree as H and we can proceed by an inductive argument; formally, we induct on m on the statement:

"Let $P(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ be a polynomial. There is some $d_0 > 0$ such that for any $d > d_0$ and any $\varepsilon > 0$, there is T_0 such that if $T > T_0$, for any subvariety $H \subseteq \mathbb{A}^m$ of degree bounded above by $P(T, d)$, the proportion of points of $\mathbb{A}^m(\overline{\mathbb{Q}})$ of height at most $\log T$ and degree at most d which belong to H is less than ε.".

Therefore, as d is big enough and T is big enough (but according to the value of d), we obtain the required result. Pulling back the proof on Y' and Y (and possibly enlarging d), this concludes the argument.

3.10 Avoiding problems in the moduli space of abelian varieties

We now describe the situation in $\mathcal{A}_g$, which is the original setting of the Masser-Zannier paper. Let us first state our result.

Theorem 3.10.1. *Let a, b, c, r be positive integers and let $\mathcal{F} = \{X_1, X_2, \ldots\}$ be a set of subvarieties of $\mathcal{A}_g$ of dimension $\leq a$, "algebro-geometric" degree $\leq b$ and field of definition of degree $\leq c$ over $\mathbb{Q}$. Let $Y \subseteq \mathcal{A}_g$ be an irreducible subvariety, not contained in any proper special subvariety of $\mathcal{A}_g$, such that for any positive integers i and n we have that $Y \cap X_i^{(n)}$ is not Zariski dense in Y. Then, for d big enough, most points of $Y(\overline{\mathbb{Q}})$ (of degree $\leq d$) are not isogenous to any point of $X_i(\overline{\mathbb{Q}})$ for any $X_i \in \mathcal{F}$.*

We point out a few remarks:

- the "algebro-geometric" degree can be defined taking an embedding over $\overline{\mathbb{Q}}$ of $\mathcal{A}_g$ (which is a quasi-projective variety) in an appropriate projective space $\mathbb{P}_m$ and using the classical degree of $\mathbb{P}_m$. Alternatively, we can dissect $\mathcal{A}_g$ into finitely many affine varieties (defined over $\overline{\mathbb{Q}}$) and take the classical degrees there. The exact choice of the degree is irrelevant, as long as such degree is comparable with the usual classical degree;

- the only substantial difference with the statement of Theorem 3.6.7 is the requirement for Y not to be contained in a proper special subvariety. This simplifies the proof substantially, since we can completely avoid dealing with (positive dimensional) weakly special subvarieties: by the arguments of Masser-Zannier combined with a result of Zywina [68], "most" points of Y are "Hodge generic"; the action of the absolute Galois group on the Tate module (of the abelian variety represented by the point) is the same, up to a finite uniform index, as the action of the absolute Galois group on the Tate module of the "generic" abelian variety of Y (i.e. the one represented by the generic point of Y). In terms of unlikely intersections, this implies by the result of Gao (Lemma 10.2.6 of an Appendix to the article by André-Corvaja-Zannier [2]) that "most" points of Y are not contained in any proper weakly special subvariety of positive dimension. Thus, all the arguments involving positive dimensional weakly special subvarieties in the previous section (namely, parts 3. and 4. of the proof) can be skipped and replaced by the familiar point counting. This is essentially the same strategy employed by Masser-Zannier in [42];

- the hypothesis of Y not contained in a proper special subvariety of the ambient space would be surprisingly less useful in $Y(1)^r$. Indeed, a "Hodge generic" point of $Y(1)^r$ is a point $(p_1, p_2, \ldots, p_r)$ where the p_i represent r non pairwise isogenous elliptic curves (none of which CM); such point is anyway contained in some positive dimensional weakly special subvarieties, for instance the "vertical" linear subspaces, none of which is special by "Hodge genericity". Thus, any two points can be "connected" by weakly special subvarieties: if $p_1, p_2 \in Y(1)^r$ then there exist two weakly special subvarieties W_1, W_2

and a "path" from p_1 to p_2 completely contained in $W_1 \cup W_2$; such "path" can be a semialgebraic curve when seen in the uniformising space obtained taking the preimages via j. Even worse, for any weakly special variety W we can "connect" p_1 to something in $W_1 \cap W$, then to something in $W_2 \cap W$ and then to p_2: we cannot even restrict our exceptions to the "vertical" linear subspaces, since arbitrary weakly special subvarieties such as W might occur. Therefore the analysis of weakly special subvarieties (part 4. of the previous section) seems inevitable in the context of $Y(1)^r$.
A consequence of the result by Gao mentioned above illustrates well the situation: a point of $\mathcal{A}_g$ is contained in a proper weakly special subvariety if and only if it is contained in a proper special variety. The analogous statement is false for $Y(1)^r$.

The proof of Theorem 3.10.1 then follows closely the proof of Masser-Zannier, plus the arguments involving degrees that allow us to deal with a family $\mathcal{F}$ of varieties rather than a single one.
We might wonder to which extent our previous proof for $Y(1)^r$ (Theorem 3.6.7) could be extended to $\mathcal{A}_g$. Let us formulate a conjectural statement:

Theorem 3.10.1 holds even if Y is contained in a proper special subvariety.

The only obstacle is the lack of an analogue of our "isogeny estimates for weakly special subvarieties" (Proposition 3.8.14); we conjecture the following analogous statement (for $\mathcal{A}_g$) to hold true.

Conjecture. *Given a positive integer g, there exists a polynomial $P(t_1, t_2, t_3, t_4) \in \mathbb{Q}[t_1, t_2, t_3, t_4]$ with the following property.*
Let $W \subseteq \mathcal{A}_g$ be a weakly special subvariety defined over $\overline{\mathbb{Q}}$ and let $p \in \mathcal{A}_g(\overline{\mathbb{Q}})$. Suppose that p is isogenous to some point of $W(\overline{\mathbb{Q}})$; then there exists a positive integer n satisfying

$$n < P(h(p), [\mathbb{Q}(p), \mathbb{Q}], \deg W, [\mathbb{Q}(W) : \mathbb{Q}])$$

such that p is isogenous to a point of $W(\overline{\mathbb{Q}})$ via an isogeny of degree n.

We will observe later that, assuming this Conjecture, we are able to get rid of the assumption of Y not being contained in a proper special variety from the statement of Theorem 3.10.1.

Before we dive into the proofs, we point out another difference between the case of $Y(1)^r$ and $\mathcal{A}_g$.
If two elliptic curves are connected by an isogeny of degree n, this gives rise to a matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with integer entries, acting on $\mathbb{H}$ and describing the isogeny. By the fundamental Lemma 2.5.12, the integers a, b, c, d can be chosen of height at most $\log n$. The entries a, b, c, d are the coordinates of the points that we are able to control with the Pila-Wilkie Theorem 2.2.5.

A similar bound cannot hold, even in principle, for $\mathcal{A}_g$: there are simple abelian varieties with infinite automorphism group (i.e. infinitely many degree one isogenies).
In the non simple case we have the example of a square of an elliptic curve E:

$$\varphi_n : E \times E \to E \times E$$

$$\varphi_n(P, Q) = (P, P + [n]Q)$$

and the morphism φ_n, for any integer n, has degree 1.
However, Masser and Zannier solved the issue completely, to the extent that we can almost forget about this by recalling their form of isogeny estimates that, instead of bounding the degree of the connecting isogeny, bound its degree together with some *length* (of the relevant Rosati quadratic form):

Theorem 3.10.2. *There exists a polynomial $P(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$, depending on g only, with the following property. Let $p, q \in \mathcal{A}_g(\overline{\mathbb{Q}})$ be isogenous. There is an isogeny between p and q whose length and degree are bounded above by*

$$P(h(p), [\mathbb{Q}(p, q) : \mathbb{Q}]).$$

Proof. The statement for the length is a combination of Lemma 4.2 and Lemma 4.4 of [42]. The statement for the degree is the usual isogeny estimates Theorem 3.2.2. □

Remark 3.10.3. *An historical remark: as Masser and Zannier point out, the original proofs for isogeny estimates by Masser and Wüstholz based on transcendental methods already implied a suitable bound on the length. In this sense, the usual isogeny estimates for bounding the degree simply "forgot" about the length in the proofs.*

Once a suitable fundamental domain (analogous to $D \subseteq \mathbb{H}$ for the case of $Y(1)$) is fixed, upper bounds on the degree and the length of an isogeny finally allow us to bound correctly the coefficients in the relevant matrices.

Remark 3.10.4. *We have not discussed at all about polarisations. Indeed, $\mathcal{A}_g$ is the coarse moduli space of principally polarised abelian varieties of dimension g. In all of our hypotheses (and in the work of Masser-Zannier as well) none of our maps, in particular our isogenies, need to respect polarisations. It is therefore not important for us to keep track of this.*

We are now ready to outline the scheme of both proofs (that we do at once). The general arguments are fairly similar either to the proofs in the paper of Masser and Zannier or to the proofs in the previous section, so we often refer to those instead of repeating the very same reasoning.

Let $p \in \mathcal{A}_g(\overline{\mathbb{Q}})$ be a point of degree d. Let $\mathcal{S}$ be a variety parameterising the subvarieties of $\mathcal{A}_g$ of dimension at most a and degree at most b. We recall that a subvariety $X \subseteq \mathcal{A}_g$ of dimension at most a and degree at most b is defined over a number field K precisely when it is represented by a point of $\mathcal{S}(K)$.

1. We construct a definable set $S_{p,X}$ parameterising "virtual isogenies" between p and any point of $X(\mathbb{C})$, requiring that integral points of $S_{p,X}$ correspond to isogenies between p and a point of $X(\mathbb{C})$. In this case, $S_{p,X}$ is a family of matrices acting on the Siegel upper half-space $\mathbb{H}_g$, that we briefly describe later. As well as before, the sets $S_{p,X}$ form a definable family as p varies in $Y(\mathbb{C})$ and X (more precisely, its representing point) varies in $\mathcal{S}(\mathbb{C})$. We also point out that an isogeny between p and some point $q \in X(\mathbb{C})$ gives rise to an integral point on $S_{p,X}$ whose height is bounded above by the length and the degree of that very isogeny.

2. We classify the semialgebraic varieties contained in $S_{p,X}$ using again the Theorem of Klingler-Ullmo-Yafaev of [34] (see Theorem 3.9.2 above) and we recover exactly the same classification as in the case of $Y(1)^r$: two points on $S_{p,X}$ are "connected" by a semialgebraic variety whenever they correspond to two points on X that lie on the same weakly special subvariety W. The converse is again false, as we can "glue' (non smoothly) two semialgebraic curves exploiting a point on $S_{p,X}$ that corresponds to a point in the intersection $W_1 \cap W_2$ for two positive dimensional weakly special subvarieties $W_1, W_2 \subseteq \mathcal{A}_g$.

3. The exact same decomposition of $S_{p,X}$ as a finite disjoint union of definable subsets that we employed before works, but we cannot rely on our explicit classification of weakly special subvarieties of $Y(1)^r$ anymore. The result of Binyamini-Daw of [9] (see also Theorem 3.9.3 above) indeed not only implies that the maximal weakly special subvariety of X have bounded degree, but it states that they are parameterised by a variety whose degree depends on $\deg X$ only. In view of a result by Grushevsky-Mondello-Salvati Manni-Tsimerman in [25], each weakly special subvariety has an analogous "special type" and we can carry out the same construction as before.
   We can therefore decompose $S_{p,X}$ according to the special types of maximal weakly special subvarieties as in the previous section. If Tp is a set of such special types, we call again $S_{p,X,Tp}$ the (definable) subset of $S_{p,X}$ consisting of points that correspond to the points of X such that the special types of the maximal weakly special subvarieties of X through them are exactly the elements of Tp. We can finally prove that whenever two points on $S_{p,X,Tp}$ are "connected" by a semialgebraic variety contained in $S_{p,X,Tp}$ then they correspond to points that lie on the same maximal weakly special subvarieties of X. As in the previous section, all the estimates here are independent of p and X (as the bound on the degree of the maximal weakly special subvarieties given by the Binyamini-Daw result depends on $\deg X$ only, which in turn is bounded by our hypotheses on $\mathcal{F}$).

4. We cannot apply our form of isogeny estimates (Proposition 3.8.14) anymore and we therefore assume Conjecture 3.10 to hold: if p is isogenous to some point on a weakly special subvariety W, then we can bound the degree of an isogeny connecting p with some point on W with a polynomial in $d, h(p), [\mathbb{Q}(W) : \mathbb{Q}]$ and $\deg W$. This latter quantity is bounded uniformly (when $X \in \mathcal{F}$ varies) by the result of Binyamini-Daw (Theorem 3.9.3), thus we can recover exactly as in the previous section that the degree of the field

of definition of W is "big" in terms of the degrees of the connecting isogenies.

We point out that if Y is not contained in a proper special subvariety then, combining the results of Gao and Zywina that we mentioned above, "most" points of $Y(\overline{\mathbb{Q}})$ of degree bounded above by any sufficiently big constant will not be contained in any positive dimensional weakly special subvariety and thus the usual isogeny estimates for abelian varieties rather than the conjectural version for weakly special subvarieties suffice.

5. We can finally apply the Pila-Wilkie Theorem on each $S_{p,X,Tp}$ exploiting the big degree of the field of definition of W, whose conjugates give rise to many integral points that must be covered by many distinct semialgebraic varieties. The upper bound on the degree of the field of definition of W obtained in this way allows us to bound the degree of the connecting isogenies.

6. We can now bound the height of the point representing X in $\mathcal{S}$. This is the very same argument as 6. in the previous section: we enlarge $\mathcal{F}$ suitably and we find a unique variety of minimal dimension containing all the Galois conjugates of q. This uniqueness gives a height bound on the representing point (in an appropriate Chow variety) of the minimal variety that we construct.

7. The counting argument of 7. above applies identically here. As long as d is big enough, the number of points of degree $\leq d$ in $Y(\overline{\mathbb{Q}})$ with height bounded by $\log T$ will be comparable to T^{d^2+d}; we just need to show that the points on $Y(\overline{\mathbb{Q}})$ with height bounded above by $\log T$ and degree $\leq d$ that lie in the intersections of the isogenous images of the minimal varieties described above are "sparse", provided that the isogenies and the (points representing the) minimal varieties satisfy the bounds we obtained before.

We proceed with outlining details for each point.

1. We refer to the construction of Masser-Zannier in [42].
   The uniformising space for $\mathcal{A}_g$ is now the Siegel upper half space $\mathbb{H}_g$, that consists of the $2g \times 2g$ symmetric matrices with complex coefficients whose imaginary part is positive definite. The symplectic group $\mathrm{Sp}_{2g}(\mathbb{R})$ acts on $\mathbb{H}_g$ and any isogeny between two abelian varieties represented by two elements of $\mathbb{H}_g$ can be represented by a symplectic matrix with integer coefficients.
   After choosing a suitable fundamental domain D, we define $S_{p,X}$ as the subset consisting of the matrices of $\mathrm{Sp}_{2g}(\mathbb{R})$ such that

   $$M \in S_{p,X} \text{ if and only if } M \text{ maps the point of } D \text{ corresponding to } p \text{ to a point of } D \text{ corresponding to any point of } X(\mathbb{C}).$$

   This definition is exactly analogous to our previous definition of $S_{p,X}$ for $Y(1)^r$.
   Integral points of $S_{p,X}$ correspond to isogenies and viceversa; moreover, Masser and

Zannier provide the following estimate which is the correct analogue of fundamental Lemma 2.5.12. We reformulate such estimate in our notation.

Lemma 3.10.5. *There is a polynomial $s(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$, depending on g only, satisfying the following property. Let $p, q \in \mathcal{A}_g$ be isogenous via an isogeny of degree n and length ℓ. There is a matrix $M \in Sp_{2g}(\mathbb{Z})$ with entries of exponential height at most $s(n, \ell)$ such that M maps the point of D corresponding to p to the point of D corresponding to q.*

This, combined with the form of isogeny estimates taking also lengths into account that we mentioned at the beginning of the section, is sufficient for the application of the Pila-Wilkie Theorem 2.2.5 that we will use later.

2. The classification of semialgebraic varieties of $S_{p,X}$ is the same as in 2. of the previous section. We quote again the hyperbolic Ax-Lindemann-Weierstrass Theorem 3.9.2 by Klingler-Yafaev-Ullmo.
We denote with π the uniformising map $\pi : \mathbb{H}_g \to \mathcal{A}_g$, suitably restricted to a fundamental domain, and, again, we extend its scope to $S_{p,X}$, by

$$\pi(M) = \pi(M \cdot \pi^{-1}(p))$$

for any $M \in S_{p,X}$.

Theorem (Hyperbolic Ax-Lindemann-Weierstrass)**.** *Let V be a semialgebraic subvariety contained in $S_{p,X}$. Any irreducible component of the Zariski closure of $\pi(V) \subseteq X$ is a weakly special subvariety and if $W \subseteq X$ is a weakly special subvariety then $\pi^{-1}(W)$ is a semialgebraic subset of $S_{p,X}$.*

This is not enough since a (non-smooth) connected semialgebraic curve in $S_{p,X}$ could account for two weakly special subvarieties W_1, W_2, exploiting the non-smoothness to "cross" from W_1 to W_2 via some point of $W_1 \cap W_2$.
We solve the issue decomposing $S_{p,X}$ exactly as in 3. above and we perform it in the next step.

3. It turns out that $S_{p,X}$ can be decomposed into the same suitable definable subsets that we defined in 3. as above; we need to check that things work out correctly without appealing to our explicit description of weakly special subvarieties anymore.
We first describe a notion of *special type* in the context of $\mathcal{A}_g$. By the classification of Grushevsky-Mondello-Salvati Manni-Tsimerman (in [25]), any weakly special subvariety W is either a point, special or a product of the form $W = H \times \{z\} \subseteq H \times H'$ where $H \times H' \subseteq \mathcal{A}_g$ is a special subvariety and $z \in H'$.
If W is a point we say that its special type is "point", while if it is of the form $H \times \{z\}$ we set H (together with the inclusion $H \times H' \subseteq \mathcal{A}_g$) to be its special type. If W is a special subvariety which is not a product as above, we set its special type equal to W itself.

The paper of Binyamini-Daw [9], quoted in 3. of the previous section, proves actually more than the bound on the degrees we have used above. Indeed, their arguments prove that all the maximal weakly special subvarieties of X arise in a finite family; moreover, this finite family can be chosen among families of a degree bounded in terms of $\deg X$ (which is uniformly bounded in $\mathcal{F}$ by hypothesis). We obtain the following.

There is a finite set TP of special types such that if $W \subseteq X_i$ is a maximal weakly special subvariety for some $X_i \in \mathcal{F}$, then the special type of W is an element of TP.

We notice that this in particular means that each point of X is contained in finitely many maximal weakly special subvarieties, since two distinct weakly special subvarieties of the same special type are disjoint.
We are now ready to perform the same construction of $S_{p,X,Tp}$ as we did before.

$$S_{p,X,Tp} = \{z \in S_{p,X} \mid \ z \text{ corresponds to a point on } X$$

$$\text{contained in exactly } |Tp| \text{ maximal weakly special subvarieties of } X,$$

$$\text{each with special type an element of } Tp\}.$$

Again, each $S_{p,X,Tp}$ is definable, since our maximal weakly special subvarieties are parameterised by a family of uniformly bounded degree, hence definable.
We can finally prove the correct statement for semialgebraic varieties for any fixed subset $Tp \subseteq TP$.

Let $q, q' \in X$ such that there is a connected semialgebraic subvariety V of $S_{p,X,Tp}$ accounting for both points (i.e. V contains two points corresponding to q and q'). Then, for any maximal weakly special subvariety $W \subseteq X$, we have that $q \in W$ if and only if $q' \in W$.

The proof is essentially the same as before, the only difference being the different notions of special types: the product of modular curves $C_1 \times C_2 \times \dots C_d$ is replaced by a special subvariety H.
First, if W is indeed special and not of the form $H \times \{z\}$ for an appropriate subvariety $H \times H' \subseteq \mathcal{A}_g$, then W is the only weakly special subvariety with such special type. The statement is then immediate, since any point of $S_{p,X,Tp}$ must correspond to a point contained in W.
Let us now assume $W = H \times \{z\}$, with H (together with $H \times H' \subseteq \mathcal{A}_g$) the special type of W, and we can construct again $\alpha : [0,1] \to S_{p,X,Tp}$, a semialgebraic curve such that $\alpha(0)$ corresponds to $q \in W$, with $W \subseteq X$ maximal weakly special.
As $\alpha(t) \in S_{p,X,Tp}$ for every $t \in [0,1]$, there are maximal weakly special subvarieties $W_t \subseteq X$ of the shape

$$H \times \{z_t\} \subseteq H \times H' \subseteq \mathcal{A}_g$$

such that $\alpha(t) \in W_t$ and $z_t \in H'$ for every $t \in [0,1]$. We again take the image of α via the uniformising map from $\mathbb{H}_g$ to $\mathcal{A}_g$, the (possibly non-algebraic) curve defined as

$$\beta : [0,1] \to H';$$

$$\beta(t) = z_t.$$

This contradicts again the Hyperbolic Ax-Lindemann-Weierstrass Theorem 3.9.2, which in this case implies that the Zariski closure of $\beta([0,1]) \subseteq H'$ is a weakly special subvariety W'. We point out that H' is a special subvariety of $\mathcal{A}_g$, indeed a Shimura variety itself; that is enough for the application of the Hyperbolic Ax-Lindemann Theorem. Again, we claim that

$$H \times W' \subseteq X.$$

Indeed, the set

$$\{P \in H' \mid H \times \{P\} \subseteq X\}$$

is Zariski closed, as it is the complement of the (open) projection on the second factor of the set

$$H \times H' \setminus ((H \times H') \cap X)$$

and again we use

$$H \times \beta([0,1]) \subseteq X$$

so the claim follows. Therefore

$$W \subseteq H \times W' \subseteq X$$

and, by maximality of W, we have

$$W = H \times W'.$$

Observe that W is just

$$H \times \beta(0)$$

i.e. $\beta : [0,1] \to H'$ is a constant curve and thus the image of $\alpha : [0,1] \to S_{p,X,Tp}$ accounts only for points of X contained in W. This proves the statement above for connected semialgebraic curves (and indeed any connected semialgebraic variety) in $S_{p,X,Tp}$.

4. Let us suppose our Conjecture 3.10 to hold; we recall the statement here.

Given a positive integer g, there exists a polynomial $P(t_1,t_2,t_3,t_4) \in \mathbb{Q}[t_1,t_2,t_3,t_4]$ with the following property.
Let $W \subseteq \mathcal{A}_g$ be a weakly special subvariety defined over $\overline{\mathbb{Q}}$ and let $p \in \mathcal{A}_g(\overline{\mathbb{Q}})$. Suppose that p is isogenous to some point of $W(\overline{\mathbb{Q}})$; then there exists a positive integer n satisfying

$$n < P(h(p), [\mathbb{Q}(p), \mathbb{Q}], \deg W, [\mathbb{Q}(W) : \mathbb{Q}])$$

such that p is isogenous to a point of $W(\overline{\mathbb{Q}})$ via an isogeny of degree n.

We notice that, as a consequence of the Binyamini-Daw results, the degree $\deg W$ of any maximal weakly special subvariety $W \subseteq X$ for any $X \in \mathcal{F}$ is bounded.
We now argue that the degree of the connecting isogeny is enough to bound also the length (although we might need to replace the isogeny - this is irrelevant for our purposes). Indeed, the degree $[\mathbb{Q}(p,q) : \mathbb{Q}]$ satisfies

$$[\mathbb{Q}(p,q) : \mathbb{Q}] \leq [\mathbb{Q}(p) : \mathbb{Q}] \cdot n^{4g^2}$$

since all the Galois conjugates of q over $\mathbb{Q}(p)$ are isogenous to p, yet they are at most the number of subgroups of $(\mathbb{Z}/n\mathbb{Z})^{2g}$ generated by n elements (as an isogeny between abelian varieties means an inclusion of lattices of index n).
We can then use the arguments by Masser and Zannier so that we bound the relevant length by a polynomial $q(t_1, t_2, t_3, t_4) \in \mathbb{Q}[t_1, t_2, t_3, t_4]$ (depending on $\mathcal{F}$ only) evaluated at $h(p), [\mathbb{Q}(p), \mathbb{Q}], \deg W, [\mathbb{Q}(W) : \mathbb{Q}]$.

Let us now assume that Y is not contained in any proper special subvariety; we follow closely the arguments by Masser-Zannier (in [42]).
As a consequence of Lemma 10.2.6 by Gao in [2] we obtain that the monodromy group of the abelian scheme given by the abelian varieties parameterised by Y is Zariski dense in $\mathrm{Sp}_{2g}(\mathbb{Z})$; this implies, by using Lemma 4.4.16 of Deligne (in [18]) as in Masser-Zannier, that the action of the absolute Galois group on some ℓ-adic Tate module of the "generic" abelian variety (represented by the generic point of Y) is a finite index closed subgroup of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$.
By a result of Zywina in [68], there is a constant c', depending on Y only, such that "most" (in our sense) points of $Y(\overline{\mathbb{Q}})$ of degree $\leq d$ (whenever d is big enough) have the following property.

The action of the absolute Galois group over the field of definition of such point on the ℓ-adic Tate module of the corresponding abelian variety is a closed subgroup of $\mathrm{GSp}_{2g}(\mathbb{Z}_\ell)$ of finite index $\leq c'$.

This implies that "most" points of $Y(\overline{\mathbb{Q}})$ are ℓ-Galois generic for some prime ℓ and therefore Hodge generic (as it is pointed out by Masser-Zannier) and, again by the result of Gao, they are not contained in any proper positive dimensional weakly special subvariety. Thus, our Conjecture 3.10 is automatically true for "most" points of $Y(\overline{\mathbb{Q}})$ whenever Y is not contained in any proper special subvariety; indeed, the case of W zero dimensional is the usual type of isogeny estimates for abelian varieties.

Again, provided p is isogenous to some point of $X(\overline{\mathbb{Q}})$ contained in a maximal weakly special subvariety $W \subseteq X$:

p is isogenous to some point of $W(\overline{\mathbb{Q}})$ via an isogeny of degree at most

$$P(h(p), [\mathbb{Q}(p) : \mathbb{Q}], [\mathbb{Q}(W, p) : \mathbb{Q}(p)])$$

for a polynomial $P(t_1, t_2, t_3) \in \mathbb{Q}[t_1, t_2, t_3]$ depending on r and $\mathcal{F}$ only.

We have replaced $[\mathbb{Q}(W) : \mathbb{Q}]$ with $[\mathbb{Q}(W, p) : \mathbb{Q}(p)]$ for convenience; this is legitimate as

$$[\mathbb{Q}(W, p) : \mathbb{Q}(p)] \leq [\mathbb{Q}(W) : \mathbb{Q}]$$

and

$$[\mathbb{Q}(W) : \mathbb{Q}] \leq [\mathbb{Q}(W, p) : \mathbb{Q}(p)][\mathbb{Q}(p) : \mathbb{Q}].$$

5. The proof of this part is pretty much the same as in the previous section, except that we need to take into account the lengths of isogenies rather than their degrees only.
Suppose that $p \in Y(\overline{\mathbb{Q}})$ is isogenous to some point of $X(\overline{\mathbb{Q}})$ lying on W for a maximal weakly special subvariety $W \subseteq X$; we notice that W can be a point (indeed, this is the case for "most" choices of $p \in Y(\overline{\mathbb{Q}})$ whenever Y is not contained in a special subvariety of $\mathcal{A}_g$, as we just argued in 4.).
We showed (assuming Conjecture 3.10) that there is some $q \in W(\overline{\mathbb{Q}})$ such that p and q are related by an isogeny of degree n and length ℓ satisfying

$$\max\{n, \ell\} < P(h(p), [\mathbb{Q}(p) : \mathbb{Q}], [\mathbb{Q}(W, p) : \mathbb{Q}(p)])$$

for a polynomial $P(t_1, t_2, t_3) \in \mathbb{Q}[t_1, t_2, t_3]$ which depends on r and $\mathcal{F}$ only. Notice that this claim holds true unconditionally if Y is not contained in a proper special subvariety (or, equivalently, by Lemma 10.2.6 of Gao [2] as quoted above, in a proper weakly special subvariety), since for W equal to a point this amounts just to the usual isogeny estimates for abelian varieties.
By our discussion in 1. of this section, a bound on the degree ($\leq n$) and the length ($\leq \ell$) of the isogeny implies that we could construct an integral point on $S_{p,X}$ whose exponential height is bounded above by $s(n, \ell)$, where $s(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ is a fixed polynomial depending on g only. The upper bound $s(n, \ell)$ is then a quantity which is bounded above polynomially in terms of $h(p), [\mathbb{Q}(p) : \mathbb{Q}]$ and $[\mathbb{Q}(W, p) : \mathbb{Q}(p)]$.
Let us consider the $[\mathbb{Q}(W, p) : \mathbb{Q}(p)]$ Galois conjugates (over $\mathbb{Q}(p)$) of W; by the result of Grushevsky-Mondello-Salvati Manni-Tsimerman quoted in 3. any weakly special subvariety of $\mathcal{A}_g$ is either special (not of "product type"), a point or of the form $H \times \{z\}$ for H special: in either case, the Galois conjugates are all disjoint (and more generally, weakly special subvarieties of the same special type are disjoint).
Exactly as in 6. of the previous section, each of these Galois conjugates gives rise to a point $q' \in W'(\overline{\mathbb{Q}})$ for some conjugate W' of W and moreover q' is related to p by an isogeny with the same degree n and length ℓ. At least

$$[\mathbb{Q}(W, p, X) : \mathbb{Q}(p, X)] \geq [\mathbb{Q}(W, p) : \mathbb{Q}(p)]/[\mathbb{Q}(X) : \mathbb{Q}]$$

many of the Galois conjugations considered above fix X and, since the degree of the field of definition of X is bounded above by our hypothesis on $\mathcal{F}$, we can find an explicit constant $e > 0$ such that there are at least $e \cdot [\mathbb{Q}(W,p) : \mathbb{Q}(p)]$ integral points on $S_{p,X}$, each accounting for different maximal special subvarieties.
Writing $S_{p,X}$ as a finite disjoint union of sets of the form $S_{p,X,Tp}$ described above in 3. of this section, we see that the $e \cdot [\mathbb{Q}(W,p) : \mathbb{Q}(p)]$ many integral points we have just produced cannot be "connected" by semialgebraic sets as they belong to disjoint maximal weakly special subvarieties and finally the Pila-Wilkie Theorem reads again:

for every $\varepsilon > 0$, there is a positive constant c such that any collection of integral points of exponential height bounded above by T accounting for different maximal weakly special subvarieties on $S_{p,X}$ is of size at most $c \cdot T^{\varepsilon}$.

This implies that, fixed a polynomial $t(t_1,t_2,t_3) \in \mathbb{Q}[t_1,t_2,t_3]$, depending on $\mathcal{F}$ only, as we described above and giving an upper bound (in terms of $h(p), [\mathbb{Q}(p) : \mathbb{Q}], [\mathbb{Q}(W,p) : \mathbb{Q}]$) on the exponential height of the integral points constructed on $S_{p,X}$ with this method, we obtain

$$e \cdot [\mathbb{Q}(W,p) : \mathbb{Q}] < c \cdot t(h(p), [\mathbb{Q}(p) : \mathbb{Q}], [\mathbb{Q}(W,p) : \mathbb{Q}])^{\varepsilon}$$

and choosing ε^{-1} bigger than the degree of $P(t_1,t_2,t_3)$ in t_3, we can finally bound the degree of the connecting isogenies in terms of p only:

$$e \cdot [\mathbb{Q}(W,p) : \mathbb{Q}] < P(h(p), [\mathbb{Q}(p) : \mathbb{Q}])$$

$$n < P(h(p), [\mathbb{Q}(p) : \mathbb{Q}])$$

where $P(t_1,t_2) \in \mathbb{Q}[t_1,t_2]$ is an appropriate polynomial depending on $\mathcal{F}$ only.
We could also bound the length, but we do not need it anymore: the use of such notion is limited to bounding the exponential height of the integral points we constructed on $S_{p,X}$ (namely, the entries of the matrices of the connecting isogenies).

6. From now on the proof is really exactly the same as in the previous section. We highlight the key steps.
As a consequence of what we have just seen, if $p \in Y(\overline{\mathbb{Q}})$ is isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$, then there is some $q' \in X(\overline{\mathbb{Q}})$ such that p and q' are related by an isogeny of degree n, with

$$n < P(h(p), [\mathbb{Q}(p) : \mathbb{Q}])$$

where $P(t_1,t_2) \in \mathbb{Q}[t_1,t_2]$ is an appropriate polynomial depending on r and $\mathcal{F}$ only.

As in 6. of the previous section, we now bound the height of X (in the sense that we bound the height of the point that corresponds to X in the variety $\mathcal{S}$ that parameterises subvarieties of $\mathcal{A}_g$ of a given degree) always in terms of $h(p)$ and $[\mathbb{Q}(p) : \mathbb{Q}]$.
We can find again an upper bound b on the "algebro-geometric" degree of any element

X_i of $\mathcal{F}$ and we define again $\mathcal{S}$ as a variety parameterising all the subvarieties of $\mathcal{A}_g$ of degree at most b, so that each X_i is represented by some element of $\mathcal{S}$. We can construct $\mathcal{V}$ using the very same strategy, defining it as

$$V \in \mathcal{V}$$

if and only if

V is a finite intersection of elements represented by $\mathcal{S}$.

As we proved in 6. of the previous section (where we never used that the relevant varieties where subvarieties of $Y(1)^r$), the set $\mathcal{V}$ is indeed a variety, since the number of intersections of elements of $\mathcal{S}$ that are required to defined an element of $\mathcal{V}$ is uniformly bounded above (and each element of $\mathcal{V}$). See 6. of the previous section for a proof of these facts.
Again, $\mathcal{V}$, when embedded in the appropriate Chow variety, is defined over $\mathbb{Q}$, because $\mathcal{S}$ is (this is the same as saying that the set of arbitrary intersection of varieties of degree at most b is stable by Galois conjugation).
Again, taking $p \in Y(\overline{\mathbb{Q}})$ isogenous to some $q \in X(\overline{\mathbb{Q}})$ with $X \in \mathcal{F}$, we set K as the field of definition of X and we notice that all the Galois conjugates of q over K, say $q_1, q_2, \dots, q_k$, lie on $X(\mathbb{C})$. These k points determine uniquely X, in the sense of 6. above: there is a minimal variety represented by an element of $\mathcal{V}$ that contains all these points (e.g. take the intersection of all the varieties represented by an element of $\mathcal{V}$ that contain $q_1, q_2, \dots, q_k$). Such minimal variety is again defined over K.
We are able to apply Lemma 3.7.8 uniformly provided we use Lemma 3.9.9 (whose proof is explicitly independent of the ambient space).

We now complete the argument as before: if $p \in Y(\overline{\mathbb{Q}})$ is isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$ defined over the number field K, there is some uniquely determined minimal variety (which is a subvariety of X) parameterised by a point of $\mathcal{V}(K)$ containing all the conjugates $q_1, q_2, \dots, q_k$. Moreover, such variety is determined by only M points among the $q_1, q_2, \dots, q_k$ by Lemma 3.9.9 and we can finally apply Lemma 3.7.8.
Noting that all the heights $h(q_1), h(q_2), \dots, h(q_k)$ are equal to $h(q)$ by Galois invariance of the Weil height, we obtain the following statement.

Let $p \in Y(\overline{\mathbb{Q}})$ be isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$ defined over the number field K. There is a point $z \in \mathcal{V}(K)$ such that q is contained in the variety represented by z and

$$h(z) < c(1 + h(q))$$

for an appropriate constant $c > 0$ depending on $\mathcal{F}$ only.

We point out again that such variety can be chosen to be a subvariety of X.

7. Also in this case, the argument is fairly similar to that in 7. of the previous section. Instead of Pazuki's result Theorem 3.5.4 we use a similar inequality by Faltings (proven in [22]). We briefly recall the main ideas.
We finally prove that "most" points of $Y(\overline{\mathbb{Q}})$ are not isogenous to any point of $X_i(\overline{\mathbb{Q}})$ for any $X_i \in \mathcal{F}$. For any positive integers $T \geq 10$ and $d \geq 2$ we define again

$$Y_{T,d} = \{p \in Y(\overline{\mathbb{Q}}) \mid h(p) < \log T \text{ and } [\mathbb{Q}(p) : \mathbb{Q}] \leq d\}.$$

We still think of d as a big fixed positive integer.
Supposing that $p \in Y_{T,d}$ is isogenous to some $q \in X$ for some $X \in \mathcal{F}$, by the conclusion of 5. of this section, there is some $q' \in X(\overline{\mathbb{Q}})$ such that p is connected to q' by an isogeny of degree n and the inequality

$$n < P(\log T, d)$$

holds for an appropriate polynomial $P(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ depending on $\mathcal{F}$ only.
Moreover, by the conclusion of 6. of this section there is a subvariety $Z \subseteq X$ such that Z is parameterised by a point z of $\mathcal{V}(K)$, where K is the field of definition of X, such that:

$$q' \in Z(\overline{\mathbb{Q}})$$
$$h(z) < c_1(1 + h(q'))$$

for an appropriate constant $c_1 > 0$ depending on $\mathcal{F}$ only.
As we mentioned before, the height of q' can be bounded by using an analogue of Pazuki's Theorem for abelian varieties (by Faltings); since q' and p are connected by an isogeny of degree n, we obtain again (implicitly using that the Faltings height is comparable to the usual Weil height)

$$h(q') < c_2(1 + h(p) + \log n) < c_3(1 + \log T + \log\log T + \log d)$$

for appropriate constants c_2, c_3 depending on $\mathcal{F}$ only (where we are using again the conclusion of 5. recalled right above). Therefore for some constant $c_4 > 0$ depending on $\mathcal{F}$ only we have

$$h(z) < c_4(1 + \log T + \log d)$$

and we recall that the degree of z over $\mathbb{Q}$ is absolutely bounded as $z \in \mathcal{V}(K)$.
We proceed again counting the number of points of $\mathcal{V}(\overline{\mathbb{Q}})$ with bounded height and bounded degree and for

$$V_{U,e} = \{z \in V(\overline{\mathbb{Q}}) \mid h(z) < \log U \text{ and } [\mathbb{Q}(z) : \mathbb{Q}] \leq e\}$$

we estimate again the size of $V_{U,c}$ for

$$U = \exp\left(c_4(1 + \log T + \log d)\right)$$

as

$$|V_{U,c}| < c_5 \cdot U^{c^2 \cdot c_6}$$

for appropriate constants c_5 and c_6 depending on $\mathcal{V}$ only (and hence on $\mathcal{F}$ only). We are using the estimates of Masser and Vaaler of [40].
Hence the set of $z \in \mathcal{V}(\overline{\mathbb{Q}})$ satisfying the two bounds

$$[\mathbb{Q}(z) : \mathbb{Q}] \leq c \text{ and } h(z) < c_4(1 + \log T + \log d)$$

has size at most

$$s(T, d)$$

for an appropriate polynomial $s(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ depending on $\mathcal{F}$ only, where c is an absolute constant implied by the statement of Theorem 3.10.1.

We recall that the points of $\mathcal{V}$ correspond to varieties with a uniformly bounded "algebro-geometric" degree, say bounded above by b', which is a quantity that depends on $\mathcal{F}$ only.
Let us summarise the observations above: if any $p \in Y_{T,d}$ is isogenous to some point of $X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$, then it is isogenous, via an isogeny of degree n, to some $q' \in X(\overline{\mathbb{Q}})$ contained into some $Z \subseteq X$ parameterised by a point of $z \in \mathcal{V}(\overline{\mathbb{Q}})$ that is chosen in a set (depending on $\mathcal{F}, T, d$ only) of size at most $s(T, d)$.
Taking the union of all these admissible varieties Z as above we obtain a (possibly reducible) variety Z' of degree at most

$$b' \cdot s(T, d)$$

so that such degree grows at most polynomially in T and d.
Also this time no isogenous image $Z'^{(n)}$ will cover Y, where given a subset $S \subseteq \mathcal{A}_g$, we denote as usual with $S^{(n)}$ the set

$$S^{(n)} = \{s \in \mathcal{A}_g \mid s \text{ is isogenous to some element of } S$$

$$\text{via an isogeny of degree } n\}.$$

The claim above holds since $Y \cap X^{(n)}$ is not Zariski dense in Y for any $X \in \mathcal{F}$ and any positive integer n; indeed Z' is a finite union of subvarieties of some $X_i \in \mathcal{F}$. It is worth pointing out that the requirement of $Y \cap X^{(n)}$ not being Zariski dense in Y rather than Y not being contained in $X^{(n)}$ is one of the few differences between the proofs we are encountering at this stage.
We can recall the bound on n, which is

$$n < P(\log T, d)$$

for an appropriate polynomial $P(t_1, t_2) \in \mathbb{Q}[t_1, t_2]$ depending on $\mathcal{F}$ only. We can bound from above the degree of $Z'^{(n)}$ by the same arguments of Masser and Zannier in a similar way as we did in the previous section. In particular, their results imply the following.

There is a polynomial $u(t) \in \mathbb{Q}[t]$ such that, if $V \subseteq \mathcal{A}_g$ is a subvariety of degree D, then $V^{(n)}$ has degree at most $D \cdot u(N)$.

We apply their result and obtain

$$\deg Z'^{(n)} \leq c_7 \cdot \deg Z \cdot u(n)$$

for some constant c_7 depending on r only and the polynomial $u(t)$ described above. Now, as a consequence of the previous estimates for n, we have that

$$\deg Z'^{(n)} \leq c_7 \cdot b' \cdot v(T,d)$$

for an appropriate polynomial $v(t_1,t_2) \in \mathbb{Q}[t_1,t_2]$ depending on $\mathcal{F}$ only.
Finally, we can sum again over the values of $n < P(\log T, d)$ and intersect with Y. Any $p \in Y_{T,d}$ which is isogenous to some $q \in X(\overline{\mathbb{Q}})$ for some $X \in \mathcal{F}$ is contained in the proper subvariety

$$Y' = Y \cap \bigcup_{n=1}^{P(\log T,d)} Z'^{(n)} \subseteq Y$$

of degree bounded above by $q(T,d)$, where $q(t_1,t_2) \in \mathbb{Q}[t_1,t_2]$ is an appropriate polynomial depending on $\mathcal{F}$ only.
We point out again that Y' might indeed depend on T and d, but its degree will grow at most like a fixed power of T and d.

The rest of the proof is exactly the same as in 7. of the previous section and now we can conclude provided d is big enough. We always reduce to the case of $Y \cong \mathbb{A}^m$, with $m = \dim Y$ and apply the same arguments as in 7. that we quickly recall: the points of $\mathbb{A}^m(\overline{\mathbb{Q}})$ of degree at most d and height at most $\log T$ are roughly T^{d^2+d}; a proper subvariety of $\mathbb{A}^m$ of degree which is (a fixed) polynomial in T and d can contain at most its degree many "vertical" hyperplanes and, once we get rid of these "few" exceptions, we can conclude the proof on each individual other hyperplane by induction on m.
This finally concludes the proof of Theorem 3.6.7.

Chapter 4

Higher dimensional avoiding problems

We now shift towards some problems which are sensibly harder in nature than finding some "avoiding points"; we try to find "avoiding curves".
We will be able to obtain much less in this case and, for simplicity, we always restrict our attention to powers of $Y(1)$.

4.1 Examples of lines avoiding points

In the following results, we use again the notion of "most", but referring to "most" lines rather than to "most" points. Indeed, the existence of the Chow variety $\mathcal{S}$ (as described in section 3.7) that parameterises lines in a given affine space allows us to see any line as a point of $\mathcal{S}$ (with the nice feature that the field of definition of the relevant line coincides with the field of definition of the corresponding point in $\mathcal{S}$).
Alternatively, we can use a more direct approach (and we often refer to this choice during computation): the height of the rational line defined as the zero locus of

$$\{(x, y) \in Y(1)^2 \mid y = ax + b\}$$

is simply the height of the point $(a, b) \in \mathbb{A}^2$. We state the first "higher-dimensional avoiding problem".

Proposition 4.1.1. *Let $(p, q) \in Y(1)^2(\overline{\mathbb{Q}})$. Then "most" rational lines "avoid" (p, q), i.e. for "most" $a, b \in \mathbb{Q}$ there is no point $(x, y) \in Y(1)^2(\overline{\mathbb{Q}})$ such that:*

1. *(x, y) is isogenous to (p, q);*
2. *$y = ax + b$.*

Proof. We give two different proofs. The first one is complete but ineffective, due to the use of the Pila-Wilkie Theorem 2.2.5. The second one is complete and effective, but based on bad reduction arguments (which would not work if a, b were assumed to be integers).
Let m, n be positive integers such that

$$\Phi_m(p, x) = 0 \text{ and } \Phi_n(q, y) = 0.$$

As in section 3.4, we can always suppose this is the case (up to switching (p, q) with (q, p)). Before diving into the proofs, we point out that the only weakly special lines (in the family considered above) are the "diagonal" line given by $a = 1, b = 0$, which is indeed special, and the "horizontal" lines given by $a = 0$. Thus, "most" lines are indeed not weakly special and we tacitly disregard weakly special lines from our discussion below.

1. The first proof uses the Pila-Wilkie Theorem 2.2.5 in the same way as we used it above many times already.
   By isogeny estimates 3.2.2 one obtains an upper bound of the shape
   $$m, n \leq c_1 \cdot [\mathbb{Q}(x, y) : \mathbb{Q}]^{2.1}$$
   for a constant c_1 depending on (p, q) only (that, we recall, is fixed once and for all).
   This allows us to bound the size of the integer coordinates of certain points on an appropriate definable set (as in section 3.4, with an identical proof) using the fundamental Lemma 2.5.12 by the very same quantity as above
   $$c_1 \cdot [\mathbb{Q}(x, y) : \mathbb{Q}]^{2.1}.$$
   We can apply the Pila-Zannier method: taking the conjugates of (x, y) over $\mathbb{Q}(p, q)$ and applying the Pila-Wilkie Theorem 2.2.5 (together with the functional transcendence Theorem 3.4.9 by Pila and Tsimerman) we obtain an absolute upper bound on the degree
   $$[\mathbb{Q}(x, y) : \mathbb{Q}] \leq c_2$$
   for a constant c_2 depending on (p, q) only. In turn, we conclude that m, n are absolutely bounded above as well and then only finitely many choices for (x, y) are possible; thus, we just need to avoid finitely many points and of course "most" rational lines avoid a given point. This proof is not effective due to the Pila-Wilkie Theorem.

2. This proof uses a bad reduction argument, revolving around the fact that if $\Phi_m(x, p) = 0$ then x and p have bad reduction at the same primes (and the same holds for y and q).
   Let r be a prime that does not appear in the "denominator" of p and q (i.e., no prime ideal lying above r appears into the prime ideal factorisation of either p or q with a negative exponent). We observe that r cannot be in the denominator of x and y as well (since isogenies preserve bad reduction at r). We refer to Proposition 7.1.9; the proof essentially amounts to observing that any modular polynomial $\Phi_n(x, y)$ has integer coefficients and it is monic in both x and y.
   Let us suppose that the rational prime r does not appear in the denominator of a and it appears in that of b, i.e.
   $$v_r(a) \geq 0 \text{ and } v_r(b) < 0$$
   for some r not appearing in the "denominators" of p and q as above; for instance, the line
   $$\{(x, y) \in Y(1)^2 \mid y = x + 1/r\}$$

has such property whenever r is big enough (in terms of (p,q)).
We can see that any of these lines avoid (p,q), since no (x,y) lying on it can have good reduction at r for both x and y.
Thus, we just need to prove the following statement, where we denote with S the (finite) set of rational primes which lie below some prime occurring in the "denominator" of p and q.

For "most" $a, b \in \mathbb{Q}$ there is a rational prime $r \notin S$ such that $v_r(a) \geq 0$ and $v_r(b) < 0$.

The proof is similar to that of Remark 3.4.4. We write $a = A/C$ and $b = B/D$ for A, B, C, D integers such that A and C are coprime and B and D are coprime as well.
Let $r \notin S$ be a fixed prime. Let $T \geq 1$ that we will later make tend to infinity; we consider the "probability" that r divides D (and hence r does not divide B) whenever a, b are chosen among all the numbers satisfying

$$\max\{h(a), h(b)\} < \log T$$

for T tending to infinity. We point out that this is exactly the same as the probability of r dividing C (and hence not dividing A).
Such "probability" is the ratio of the following values (the first one obtained already in Remark 3.4.4):

$$\sum_{i \leq T} \varphi(i) = \frac{6}{\pi^2} T^2 + C_1 \text{ with } |C_1| \leq 3T^{5/3}$$

$$\sum_{\substack{i \leq T \\ r|i}} \varphi(i) = \sum_{\substack{i \leq T \\ r^2|i}} \varphi(i) + \sum_{\substack{i \leq T \\ r|i \\ r^2 \nmid i}} \varphi(i) = r \sum_{\substack{i \leq T/r \\ r|i}} \varphi(i) + (r-1) \sum_{\substack{i \leq T/r \\ r \nmid i}} \varphi(i) =$$

$$= \sum_{\substack{i \leq T/r \\ r|i}} \varphi(i) + (r-1) \sum_{i \leq T/r} \varphi(i) = \sum_{\substack{i \leq T/r \\ r|i}} \varphi(i) + \frac{r-1}{r^2} \cdot \frac{6}{\pi^2} T^2 + C_2 \text{ with } |C_2| \leq 3T^{5/3}$$

from which we deduce that the "probability" that D is zero modulo r is

$$\frac{r-1}{r^2} \left(1 + \frac{1}{r^2} + \frac{1}{r^4} + \dots\right) = \frac{1}{r+1}$$

and therefore the "probability" that the property above is satisfied, namely that D is zero modulo r and B is nonzero modulo r, tends to

$$\frac{r}{(r+1)^2}$$

when T tends to infinity.
As these probabilities tend to be independent when combined with different primes, the fact that the series of the reciprocal of the primes diverges immediately implies that for most choices of a and b there will be a prime r satisfying the property above.
This completes the proof, which is also effective; the main issue here is that this argument fails whenever any integrality condition on a and b are required.

□

In general it is very hard to find lines that "avoid" given positive dimensional varieties or even countable families of points (the result above being that of a single point). In section 4.2 we describe a result, Theorem 4.2.1, which shows that "most" lines "avoid" a curve, provided it is "highly asymmetrical".
For instance, we do not know if we can "avoid" all the points of $Y(1)^2(\mathbb{Q})$.

Question. *Does there exist a curve $C \subseteq Y(1)^2$ over $\overline{\mathbb{Q}}$ such that there are no point $(p,q) \in Y(1)^2(\mathbb{Q})$ and point $(x,y) \in C(\overline{\mathbb{Q}})$ such that (p,q) and (x,y) are isogeneous?*

Remark 4.1.2. *A possible approach could come from the "solvability" (conjectural) arguments in section 7.3.*

This question is analogous to a result by Pila on a modular Fermat equation (in [50]).

We can give a partial answer if we restrict our set "to avoid" to the parabola $\{(q,q^2) \mid q \in \mathbb{Q}\}$.

Proposition 4.1.3. *Let K be a number field of degree 4. For "most" lines $\ell \subseteq Y(1)^2$ defined over K there are no points $(q,q^2) \in Y(1)^2(\mathbb{Q})$ and $(x,y) \in \ell(\overline{\mathbb{Q}})$ such that (x,y) and (q,q^2) are isogenous.*

The proof is fairly similar to the argument using isogeny estimates Theorem 3.2.2 and the Pila-Wilkie Theorem 2.2.5 recalled at the beginning of this section, together with the Theorem of Pazuki 3.5.4, stating that isogenous moduli have comparable heights (in terms of the degree of the connecting isogeny), which is necessary in order to exploit the "asymmetry" of the point (q,q^2).

Proof. Whenever the line ℓ is defined by

$$\{(x,y) \in Y(1)^2 \mid y = ax + b\}$$

with $a, b \in K$, we denote with $h(\ell)$ the quantity $\max\{h(a), h(b)\}$.
Let $(q,q^2) \in Y(1)^2(\mathbb{Q})$ and $(x,y) \in \ell(\overline{\mathbb{Q}})$ be related by isogenies of degree m, n:

$$\Phi_m(q,x) = 0 \text{ and } \Phi_n(q^2,y) = 0.$$

Applying isogeny estimates together with the Pila-Wilkie Theorem, as before, we obtain:

$$[\mathbb{Q}(x,y) : \mathbb{Q}] \leq c_1 h(q)$$

$$\max\{m,n\} \leq c_2 (h(q))^5$$

for two absolute, but ineffective, constants c_1, c_2; we point to section 3.4 for all the details.
We now use Pazuki's Theorem 3.5.4 together with the elementary height properties of Proposition 2.6.7 and we obtain the following three inequalities:

$$|h(q) - h(x)| \leq 60 \log h(q) + c_3$$

$$|h(x) - h(y)| \leq 2h(\ell) + 2$$
$$\left|h(y) - h(q^2)\right| \leq 60 \log h(q) + c_3$$

which, combined, give

$$h(q) = \left|h(q) - h(q^2)\right| \leq 120 \log h(q) + 2h(\ell) + 2c_3$$

where c_3 is an absolute constant. We point out that this very last step (i.e. $h(q^2) = 2h(q)$) is where the "asymmetry" comes into play.
We conclude that

$$h(q) \leq 3h(\ell) + c_4$$

for a constant c_4.
Let us now fix some $T \geq 2$ (that we think of as tending to infinity) and let us also fix some $a \in K$ with $h(a) < \log T$. We allow $b \in K$ to vary as long as $h(b) < \log T$ and we thus produce around T^8, since $[K : \mathbb{Q}] = 4$ many disjoint lines; we are exploiting the estimates of Masser and Vaaler, as in [40].
Any candidate (q, q^2) isogenous to any point on the lines just constructed above must satisfy $h(q) < 3 \log T + c_4$ and, since $q \in \mathbb{Q}$, there are roughly T^6 choices for these. As m, n are then bounded by

$$c_5 (\log T)^5$$

for some constant c_5, we finally obtain that the isogenous images of the admissible (q, q^2) are at most

$$c_6 \cdot T^{6.1}$$

for some constant c_6. We recall that a is fixed and thus any of the points produced above can account for at most one of these lines: this completes the proof.

□

4.2 Lines avoiding curves

We now investigate a broader class of "avoiding problems" inspired by the result of Masser-Zannier.
We have mostly been concerned with finding many points with(out) a certain property, namely "avoiding" a variety (or even a countable family of varieties).
In this section we wish to find positive dimensional varieties such that *all* of their points have such an "avoiding property". Taking into account the usual unlikely intersections reasoning concerning dimensions, we can summarise our principle as follows.

Given a Shimura variety S and a proper subvariety $X \subseteq S$, then "most" subvarieties $Y \subseteq S$ of codimension $1 + \dim X$ have the following property: there are no points $p \in X(\overline{\mathbb{Q}})$, $q \in Y(\overline{\mathbb{Q}})$ such that p and q are isogenous.

Here "most" can be made meaningful using the notion of Chow variety that we have used already in section 3.7: subvarieties of S of fixed dimension and degree can be parameterised by a variety (indeed, a Chow variety) and therefore, once the degree of Y is fixed, we can "count" the varieties and we can therefore use a sensible notion of "most".
In practice, in this case our examples will be much more explicit and there will be no need for abstract Chow varieties.
Indeed, the principle outlined above is very far from being proven. For instance, we do not even know whether the following holds.

Question. *Let $C \subseteq Y(1)^3$ be a curve over $\overline{\mathbb{Q}}$. Is it true that for "most" rational lines ℓ there are no points $(\alpha_1, \alpha_2, \alpha_3) \in C(\overline{\mathbb{Q}})$ and $(\beta_1, \beta_2, \beta_3) \in \ell(\overline{\mathbb{Q}})$ such that $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ are isogenous?*

Let us fix a way of parameterising the lines in $Y(1)^3$; we point out that it is not in principle obvious that an "avoiding problem" holding for a certain parameterisation of lines should hold with another, although in practice the exact parameterisation does not seem to matter.
Let us then choose to parameterise our lines by $\mathbb{A}^4$, with coordinates (a, b, c, d), in the sense that the point $(a, b, c, d) \in \mathbb{A}^4$ corresponds to the line

$$\{(t, at + b, ct + d) \in Y(1)^3 \mid t \in Y(1)\}.$$

Although we are missing some lines (namely, the lines whose first coordinate is constant), these are contained in a proper closed subvariety of the Chow variety parameterising all the lines and we might forget about them while dealing with density (i.e. "most") problems.

We can give an affirmative answer to the question above provided C is assumed to be highly asymmetric. We denote with x_1, x_2, x_3 the coordinates of each factor of $Y(1)^3$.

Theorem 4.2.1. *Let $C \subseteq Y(1)^3$ be a curve over $\overline{\mathbb{Q}}$ satisfying*

$$0 < 10^4 \deg_{x_1}(C) < 10^2 \deg_{x_2}(C) < \deg_{x_3}(C).$$

For "most" rational lines ℓ defined as

$$\{(t, at + b, ct + d) \in Y(1)^3 \mid t \in Y(1)\}$$

there are no positive integers m_1, m_2, m_3, any point $(\alpha_1, \alpha_2, \alpha_3) \in C(\overline{\mathbb{Q}})$ and any point $(\beta_1, \beta_2, \beta_3) \in \ell(\overline{\mathbb{Q}})$ such that

$$\Phi_{m_i}(\alpha_i, \beta_i) = 0 \text{ for } i = 1, 2, 3$$

i.e. with $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ being isogenous.

Remark 4.2.2. *We point out that such an asymmetric curve as above can never be contained in any proper weakly special subvariety of $Y(1)^3$ and this simplifies our subsequent application of the Pila-Wilkie Theorem 2.2.5.*

This fact follows from our explicit classification of the weakly special subvarieties of $Y(1)^3$, as described in section 3.8. Indeed, any proper weakly special subvariety which is not special is constant along at least one coordinate, while any proper special subvariety which is nonconstant along every coordinate needs to satisfy identically a relation of the type

$$\Phi_n(x_i, x_j) = 0$$

for some positive integer n and some $1 \leq i < j \leq 3$. This requirement contradicts the asymmetry of the degrees by the following argument exploiting heights.
By Pazuki's Theorem 3.5.4, the heights $h(x_i)$ and $h(x_j)$ have a bounded difference for every point on the curve (since the degree of the connecting isogeny is prescribed); viceversa, since $\deg_{x_i}(C) \neq \deg_{x_j}(C)$, the ratio

$$h(x_j)/h(x_i) \text{ tends to } \deg_{x_j}(C)/\deg_{x_i}(C)$$

by a result of Néron in [46] *that we also recall later. This is seen to be false by taking any point on the curve with sufficiently big height.*

We postpone for a bit the proof of this Theorem, since we will need the following statement.

Proposition 4.2.3. *For "most" $a, b \in \mathbb{Q}$, the following inequality holds:*

$$\inf_{t \in \overline{\mathbb{Q}}} \max\{h(t), h(at+b)\} > \frac{1}{7}\max\{h(a), h(b)\}.$$

Remark 4.2.4. *This Proposition means that most rational lines contain no point whose height is "small" compared to that of the line. For instance, a counterexample is given by the family of lines of $Y(1)^2$*

$$\{(x, y) \mid y = mx - m\}$$

for any rational m; these lines always contain the point $(1, 0)$, even if m has a "big" height.
As we have already observed above, a line that contains two points with "small" height needs to have a "small" height as well, since we can recover the coefficients of the line from the two points; we recall our very general interpolation result, Lemma 3.7.8.

Proposition 4.2.3 allows us to control the height of at least one coordinate of any point of ℓ, for most choices of the line: it is possible that some line with "big" height has a point with "small" height (which turns out to be hard to control), but we can show that this does not happen for most lines. We then use Pazuki's Theorem 3.5.4 in order to contradict our asymmetry assumption, as in Proposition 4.1.3.

Proof. We prove Proposition 4.2.3 in two steps, according to whether $t \in \mathbb{Q}$ or not.

1. We suppose that $t \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$; in this case the proof is rather easy, as we can recover the coefficients a, b of the line by taking Galois conjugates.
   Let us assume we have two "exceptions" t_1, t_2:

$$h(t_1), h(at_1+b), h(t_2), h(at_2+b) < \frac{1}{7}\max\{h(a), h(b)\}.$$

If we denote with $u_1 = at_1 + b$, $u_2 = at_2 + b$ and $f = \frac{1}{7}\max\{h(a), h(b)\}$, we have

$$a = \frac{u_2 - u_1}{t_2 - t_1} \text{ and } b = u_1 - at_1$$

and, by the elementary properties of the height given by Proposition 2.6.7,

$$h(a) < 4f + 3 \text{ and } h(b) < 6f + 5$$

from which we obtain $h(a), h(b) < 35$ (and we conclude).

Remark 4.2.5. *A similar strategy works for an arbitrary curve E if we wish to obtain some information on the "essential infimum", which is defined as*

$$\sup_{U \subseteq E} \left(\inf_{P \in U(\overline{\mathbb{Q}})} h(P) \right)$$

where $U \subseteq E$ ranges among the cofinite subsets of E. Thus, the essential infimum of a curve is comparable with its height (in our usual sense, i.e. the height of the corresponding point in a Chow variety).

2. We now deal with the case $t \in \mathbb{Q}$, which requires more work.
We write $a = A/C$, $b = B/D$ and $t = X/Y$ as irreducible fractions. The integers A, B, C, D vary among integers with height $\leq \log T$ and X, Y are integers with height $\leq (\log T)/7$. The key step is proving that

$$\gcd(ADX + BCY,\ CDY)$$

(which corresponds to the "cancellation" when computing $at+b$) satisfies a suitable upper bound (for every X, Y with small height) for most possibilities of A, B, C, D.
We can assume for simplicity A and B (and respectively C and D) to be pairwise coprime, since in most cases $\gcd(A, B)$ (and $\gcd(C, D)$) will be very small.
For instance, the choices of A, B, C, D as above are roughly

$$\frac{36}{\pi^4} T^4$$

as we have observed in section 3.4. If we impose the additional conditions

$$u \mid \gcd(A, B) \text{ and } v \mid \gcd(C, D)$$

then the number of choices of A, B, C, D as above are then of a size comparable to

$$\frac{36}{\pi^4} \frac{T^4}{u^2 v^2}$$

and it can then be concluded using an "inclusion-exclusion" argument that "small" gcds will make up the vast majority.
Let us take a prime number p and a positive integer k; we suppose that

$$p^k \mid ADX + BCY \text{ and } p^k \mid CDY.$$

We have three possibilities:

(a) $v_p(C) \geq k/3$: then by the coprimality assumption $v_p(X) \geq k/3$;

(b) $v_p(D) \geq k/3$: then by the coprimality assumption $v_p(Y) \geq k/3$;

(c) $v_p(Y) \geq k/3$.

In any case,

$$k \leq 3(v_p(X) + v_p(Y))$$

holds. Putting together these inequalities for all the possible prime powers p^k, we get

$$h(\gcd(ADX + BCY,\ CDY)) \leq 3(h(X) + h(Y)) \leq (6/7)\log T.$$

In contrast, for most values of C, D, we obtain

$$h(CDY) \geq (13/7)\log T$$

since C and D will mostly have a height which is very close to $\log T$. This proves that

$$\frac{ADX + BCY}{CDY} = at + b$$

has height at least $\log T$ and we conclude.

□

We can now prove Theorem 4.2.1 in its entirety.
For a rational line ℓ defined as

$$\{(t, at + b, ct + d) \in Y(1)^3 \mid t \in Y(1)\}$$

we denote with $h(\ell)$ the quantity

$$\max\{h(a), h(b), h(c), h(d)\}.$$

Let us take $(\alpha_1, \alpha_2, \alpha_3) \in C(\overline{\mathbb{Q}})$ and $(\beta_1, \beta_2, \beta_3) \in \ell(\overline{\mathbb{Q}})$, that we suppose being isogenous via isogenies of degree m_1, m_2, m_3, i.e.

$$\Phi_{m_i}(\alpha_i, \beta_i) = 0 \text{ for } i = 1, 2, 3.$$

Let us denote $h = \max\{h(\beta_1), h(\beta_2), h(\beta_3)\}$, $d = [\mathbb{Q}(\beta_1, \beta_2, \beta_3) : \mathbb{Q}]$, $d' = [\mathbb{Q}(\alpha_1, \alpha_2, \alpha_3) : \mathbb{Q}]$ and $m = \max\{m_1, m_2, m_3\}$.

We first sketch the structure of the proof.

1. We would wish to apply isogeny estimates (Theorem 3.2.2) combined with the Pila-Wilkie Theorem 2.2.5 in order to bound d, d' in terms of h only (as we are used to), but the situation is slightly different from those we have encountered above since in this case both the points $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ are allowed to vary while taking Galois conjugates and, in principle, distinct pairs of points, $(\alpha_1, \alpha_2, \alpha_3)$ isogenous to $(\beta_1, \beta_2, \beta_3)$ and $(\alpha_1', \alpha_2', \alpha_3')$ isogenous to $(\beta_1', \beta_2', \beta_3')$, could give rise to the same integral point on the relevant definable set. The issue is sorted with an o-minimality argument that enables us to bound uniformly the number of pairs corresponding to the same integral point.

2. After the functional transcendence arguments are completed, we finally apply isogeny estimates and the Pila-Wilkie Theorem, bounding both d, d' and m in terms of h only.

3. Once both the algebraic degrees d, d' and the degree of the connecting isogenies m are bounded in terms of h, we apply our Proposition 4.2.3 together with Pazuki's Theorem 3.5.4 in order to exploit the asymmetry condition of the curve and concluding the proof.

Let us then dive into the details.

1. Let us consider the family of definable sets

$$S_{a,b,c,d} = \{(A_1, B_1, C_1, D_1, A_2, B_2, C_2, D_2, A_3, B_3, C_3, D_3) \in \mathbb{R}^8 \mid$$

$$\text{there exist } (x_1, x_2, x_3) \in C(\mathbb{C}) \text{ and } (y_1, y_2, y_3) \in \ell(\mathbb{C}) \text{ with}$$

$$\frac{A_i j^{-1}(x_i) + B_i}{C_i j^{-1}(x_i) + D_i} = j^{-1}(y_i) \text{ for } i = 1, 2, 3\}$$

which is an o-minimal family parameterised by (a, b, c, d); the domain for j is restricted as usual to the fundamental domain $D \subseteq \mathbb{H}$.
We observe that whenever the entries $A_1, B_1, C_1, D_1, A_2, B_2, C_2, D_2, A_3, B_3, C_3, D_3$ are integers they describe an isogeny between some $(x_1, x_2, x_3) \in C(\mathbb{C})$ and some $(y_1, y_2, y_3) \in \ell(\mathbb{C})$ (and viceversa).

We have observed in Remark 4.2.2 that C is not contained in any proper weakly special subvariety of $Y(1)^3$ and we claim that for most choices of (a, b, c, d) the same holds for ℓ: indeed, referring to the explicit description of section 3.8, as long as a and c are nonzero, then no coordinate is constant along ℓ and we observe that the line given by the equation

$$\{(x, y) \mid y = ax + b\} \subseteq Y(1)^2$$

is not contained in any modular curve (i.e. a special subvariety of $Y(1)^2$) unless $a = 1$ and $b = 0$; this is an immediate consequence of the fact that the modular curve Y_n is an irreducible curve of "geometric degree" (in the sense of section 3.7) $2\psi(n)$. The same holds for c and d.

We now need to classify the semialgebraic varieties contained in $S_{a,b,c,d}$; this is performed with a slightly different method from the previous ones, because a fixed choice of the entries A_i, B_i, C_i, D_i does not determine uniquely x_i and y_i.
Let us then suppose that a semialgebraic curve is contained in $S_{a,b,c,d}$, i.e. there are algebraic functions

$$A_i, B_i, C_i, D_i : I \to \mathbb{R} \text{ for } i = 1, 2, 3$$

where $I \subseteq \mathbb{R}$ is an open interval together with *analytic* functions

$$x_i, y_i : I \to \mathbb{R} \text{ for } i = 1, 2, 3$$

satisfying

$$\frac{A_i(t)j^{-1}(x_i(t)) + B_i(t)}{C_i(t)j^{-1}(x_i(t)) + D_i(t)} = j^{-1}(y_i(t)) \text{ for } i = 1, 2, 3\}.$$

It might not be in principle obvious to see that $x_i(t)$ and $y_i(t)$ can be made to satisfy an analytic (or even continuous) dependence on t, but this is a consequence of one of the many nice properties of o-minimal structures: we refer to the so-called "Definable Choice" property of o-minimal structures, which states that every surjective o-minimal function admits an o-minimal section. By shrinking I suitably, such section can even be taken analytic by the Analytic Cell Decomposition Theorem for $\mathbb{R}_{an,exp}$. We refer to the notes by Starchenko [62]. Let us set

$$u_i(t) = j^{-1}(x_i(t)) \text{ and } v_i(t) = j^{-1}(y_i(t))$$

that are then analytic function defined on an open interval $I \subseteq \mathbb{R}$. We can enlarge I to a complex open subset U (provided we previously shrank I appropriately in order for our algebraic functions not to be ramified) and all the relevant functions are then extended to holomorphic functions on U. It could be necessary to enlarge the (non open) fundamental domain D a bit, as usual; this is not a serious threat to our proof.
We can finally apply the Ax-Schanuel Theorem 3.4.9 for the j-function, as proven by Pila-Tsimerman in [54] and therefore the field

$$\mathbb{C}(t, u_1(t), u_2(t), u_3(t), v_1(t), v_2(t), v_3(t), x_1(t), x_2(t), x_3(t), y_1(t), y_2(t), y_3(t))$$

has transcendence degree at least 6 over $\mathbb{C}(t)$, unless "something weakly special happens", in a sense that we make precise below.
First, we point out that the relevant transcendence degree is at most 5, since the coordinates $x_1(t), x_2(t), x_3(t)$ (and likewise $y_1(t), y_2(t), y_3(t)$) combined contribute with at most 1 to the transcendence degree, as they are coordinates of points on an algebraic curve in $Y(1)^3$, and for each $1 \leq i \leq 3$, the two holomorphic functions $u_i(t)$ and $v_i(t)$ are algebraically dependent over $\mathbb{C}(t)$ by the assumption on $A_i(t), B_i(t), C_i(t), D_i(T)$ being algebraic functions of t. Thus, something "weakly special" needs to happen.
Namely, either:

- *all* the coordinates $x_1(t), x_2(t), x_3(t)$ (or $y_1(t), y_2(t), y_3(t)$) are constant: this boils down to the usual argument that we have seen for instance in section 3.4. As long as neither $\mathbb{C}$ nor ℓ are not contained in a proper weakly special subvariety of $Y(1)^3$, as it is for "most" choices of a, b, c, d, as we observed above, a semialgebraic curve will account only for a fixed point of $\mathbb{C} \times \ell$;
- some among the $x_i(t)$ or $y_i(t)$ (for any $1 \leq i \leq 3$) are constant: then (since not all of the $x_i(t)$ or $y_i(t)$ are constant) one of these coordinates needs to be constant along the whole of C or ℓ. This is not the case since "usually" neither C nor ℓ are contained in a proper weakly special subvariety;

- there is a "special relation"

$$\Phi_n(x_i(t), x_k(t)) = 0$$

occurring identically for some $1 \leq i < k \leq 3$ (or likewise for $y_i(t)$ and $y_k(t)$) and a positive integer n. We can now assume that none of the coordinates is constant, as we have treated this case right above, and thus the whole of C or ℓ is contained in a proper special subvariety of $Y(1)^3$. This does not happen "generally";

- there is a "special relation" among $x_i(t)$ and $y_k(t)$ for some $1 \leq i, k \leq 3$; we assume without loss of generality k being equal to 1:

$$\Phi_n(x_i(t), y_1(t)) = 0.$$

We now apply the Ax-Schanuel Theorem for the j-function (Theorem 3.4.9) to a restricted subset of coordinates:

$$\mathbb{C}(t, u_2(t), u_3(t), v_2(t), v_3(t), x_2(t), x_3(t), y_2(t), y_3(t))$$

by our newly added hypothesis, all the coordinates $x_1(t), x_2(t), x_3(t), y_1(t), y_2(t), y_3(t)$ are algebraically dependent; $u_2(t)$ and $v_2(t)$ are algebraically dependent by the definition of $S_{a,b,c,d}$ (and the same holds for $u_3(t)$ and $v_3(t)$) and therefore the relevant transcendence degree is at most 3.
Thus, something "weakly special" must happen; as we have already excluded all the alternatives, we are only left with the condition

$$\Phi_{n'}(x_r(t), y_s(t)) = 0$$

being identically satisfied on U for a positive integer n' and some $2 \leq r, s \leq 3$.
We argue that the two conditions

$$\begin{cases} \Phi_n(x_i(t), y_1(t)) = 0 \\ \Phi_{n'}(x_r(t), y_s(t)) = 0 \end{cases}$$

do not hold for "most" choices of (a, b, c, d). First, if $i = r$, then the line ℓ needs to be contained in a special subvariety as some relation of the form

$$\Phi_N(y_1(t), y_s(t)) = 0$$

is satisfied identically on U (for nonconstant coordinates $y_1(t)$ and $y_s(t)$) for some appropriate positive integer N.
Otherwise, if $i \neq r$, we suppose (by permuting the coordinates) the system above to be of the shape

$$\begin{cases} \Phi_n(x_1(t), y_1(t)) = 0 \\ \Phi_{n'}(x_2(t), y_2(t)) = 0 \end{cases}$$

and hence the curve C' given by the projection of C onto the first two coordinates and the line ℓ' defined by

$$\{(x, y) \in Y(1)^2 \mid y = ax + b\}$$

are "isogenous curves" in the language of section 7.1, in the sense that all the points of ℓ' are related to some point of C' by a fixed isogeny.
We prove indeed in section 7.1, using o-minimality (or even with effective arguments), that there are only finitely many choices of a and b for which C' and ℓ' can be "identically isogenous" and therefore this exception holds "sparsely".

We therefore see that the only "weakly special exception" that could occur generally is the first one, which allows anyway to conclude (using the usual argument of section 3.4, i.e. with a fixed point rather than a pair of "moving" points) that $S_{a,b,c,d}$ does not contain any positive dimensional connected semialgebraic variety accounting for more than one point. Thus, any connected semialgebraic variety contained in $S_{a,b,c,d}$ accounts for just one point of $C \times \ell$.

Before we can finally use the Pila-Wilkie Theorem 2.2.5, we need to overcome an additional obstacle. Indeed, a choice of integral entries A_i, B_i, C_i, D_i for every $1 \leq i \leq 3$ might correspond to more than one point of $C \times \ell$, i.e. to more than one choice of a point on C is isogenous to a point on ℓ.
Luckily, the number of different points a single choice of integral entries can account for is finite and uniformly bounded. Indeed, the family

$$\mathcal{T} = \{(x_1, x_2, x_3) \in C(\mathbb{C}) \text{ and } (y_1, y_2, y_3) \in \ell(\mathbb{C}) \text{ such that}$$

$$\frac{A_i j^{-1}(x_i) + B_i}{C_i j^{-1}(x_i) + D_i} = j^{-1}(y_i) \text{ for } i = 1, 2, 3\}$$

is a definable family (in the parameters A_i, B_i, C_i, D_i and also in the underlying a, b, c, d). By the property of o-minimal structures, for whichever choice of parameters $\mathcal{T}$ is finite then its size is bounded independently of the parameters.
If for a choice of integer entries A_i, B_i, C_i, D_i for every $1 \leq i \leq 3$ we had an infinite $\mathcal{T}$, then C and ℓ would be "isogenous curves" (since fixing the matrix representing an isogeny gives an explicit description of its degree, i.e. we are fixing an algebraic correspondence). We have observed already that "usually" (in terms of the choice of a, b, c, d) the line ℓ is "unrelated" to C and thus $\mathcal{T}$ will be finite (and uniformly bounded by a constant f).

Remark 4.2.6. *We point out that these "exceptional" lines ℓ that are "identically isogenous" to C cannot be in principle avoided: C can be chosen as the "isogenous image" of some line ℓ and thus all the points on ℓ will be isogenous to some point on C.*

2. Let us recall the notation introduced previously: $h = \max\{h(\beta_1), h(\beta_2), h(\beta_3)\}$, $d = [\mathbb{Q}(\beta_1, \beta_2, \beta_3) : \mathbb{Q}]$, $d' = [\mathbb{Q}(\alpha_1, \alpha_2, \alpha_3) : \mathbb{Q}]$ and $m = \max\{m_1, m_2, m_3\}$.

   We can finally really use the Pila-Wilkie Theorem in order to bound d and d' in terms of h. Indeed, if K is the field of definition of C, we obtain at least

   $$\max\{d, d'\}/[K : \mathbb{Q}]$$

   pairs of isogenous points on C and ℓ, which then correspond to

   $$\max\{d, d'\}/(f \cdot [K : \mathbb{Q}])$$

   integral points on $S_{a,b,c,d}$ with exponential height at most m (as usual, we are appealing to the fundamental Lemma 2.5.12). We recall that f is the uniform upper bound for the number of points corresponding to the same integral entries.
   The isogeny estimates Theorem 3.2.2 gives us the upper bound for m

   $$m \leq c_1 \cdot h^{2.1}(d \cdot d')^{2.1}$$

   where c_1 is a universal explicit constant. Combining the two estimates together and using the Pila-Wilkie Theorem 2.2.5, we obtain

   $$\max\{d, d'\} \leq c_2 \cdot h$$

   $$m \leq c_2 \cdot h^6$$

   for an absolute (ineffective) constant c_2 depending on C only.

3. We observe that

   $$|h(\beta_i) - h| < 4h(\ell) + 10$$

   for any $1 \leq i \leq 3$, where $h(\ell) = \max\{h(a), h(b), h(c), h(d)\}$.

   We can now finally use our Proposition 4.2.3. For simplicity, we can even assume all our values $h(a), h(b), h(c), h(d)$ to be very close to $h(\ell)$, for instance requiring them to be greater than $(1 - 1/8)h(\ell)$; this indeed happens for "most" choices of ℓ.
   Thus, in most cases

   $$h(\beta_i) > \frac{1}{8}h(\ell)$$

   for some $1 \leq i \leq 3$; we point out that Proposition 4.2.3 guarantees the inequality above to hold for *every* point $(\beta_1, \beta_2, \beta_3) \in \ell(\overline{\mathbb{Q}})$. This gives

   $$h(\ell) < 8h(\beta_i) \leq 8h.$$

   Let us choose some $1 \leq j \leq 3$ different from i. We suppose $i < j$ (the other case being analogous). We are now ready to apply Pazuki's Theorem 3.5.4:

   $$|h(\alpha_i) - h(\beta_i)| \leq 12 \log m \leq c_3 + 72 \log h$$

$$|h(\alpha_j) - h(\beta_j)| \leq 12 \log m \leq c_3 + 72 \log h$$

for a constant c_3 depending on C only; by the elementary height properties of Proposition 2.6.7

$$|h(\beta_i) - h(\beta_j)| \leq 4h(\ell) + 10 \leq 32h + 10.$$

Combining the last three equations, we get

$$|h(\alpha_i) - h(\alpha_j)| < 2c_3 + 32h + 144 \log h < c_4 + 33h$$

for a suitable constant c_4 depending on C only.
We can now use that the values of the heights of the coordinates are roughly proportional to their degrees, a classical result by Néron, [46]. This reads as

$$\left|\deg_{x_j}(C)h(\alpha_i) - \deg_{x_i}(C)h(\alpha_j)\right| < c_5 \cdot \sqrt{h(\alpha_i)} < c_6 \cdot \sqrt{1+h}$$

for some constants c_5, c_6 depending on C only. By the triangle inequality

$$\left|\left(1 - \frac{\deg_{x_i}(C)}{\deg_{x_j}(C)}\right) h(\alpha_i)\right| \leq \frac{\deg_{x_i}(C)}{\deg_{x_j}(C)} |h(\alpha_i) - h(\alpha_j)| + \left|h(\alpha_i) - \frac{\deg_{x_i}(C)}{\deg_{x_j}(C)} h(\alpha_j)\right|$$

and combining the last inequalites we get

$$0.99h(\alpha_i) \leq \left|\left(1 - \frac{\deg_{x_i}(C)}{\deg_{x_j}(C)}\right) h(\alpha_i)\right| < \frac{c_4 + 33h}{100} + c_7 \cdot \sqrt{1+h}$$

for a constant c_7 depending on C only. We observe that Pazuki's Theorem also implies (as long as h is sufficiently big)

$$(8/9)h < h - 72 \log h - c_3 \leq h(\beta_i) - 72 \log h - c_3 \leq h(\alpha_i)$$

and therefore

$$0.88h < 0.34h + 0.01c_4 + c_7 \cdot \sqrt{1+h}$$

which is a contradiction, as long as $h(\ell)$ big enough, for most choices of ℓ. We point out that Proposition 4.2.3 is crucially exploited as we just claimed that $h(\ell)$ tending to infinity implies h tending to infinity.
This completes the proof.

4.3 Curves avoiding curves

In this section we show that finding at least one curve "avoiding" a given one is easier than proving "avoidance" for most curves.
We have the following result.

Proposition 4.3.1. *Let $C \subseteq Y(1)^3$ be a curve defined over $\overline{\mathbb{Q}}$; there is a curve E that "avoids" C.*
Namely, if C is not contained in any proper weakly special subvariety of $Y(1)^3$, there are a positive integer M and an integer q such that the curve E_q defined as

$$\{(t+q, t, t^M) \in Y(1)^3 \mid t \in Y(1)\}$$

avoids C, i.e. there are no positive integers m_1, m_2, m_3 and points $(\alpha_1, \alpha_2, \alpha_3) \in C(\overline{\mathbb{Q}})$ and $(\beta_1, \beta_2, \beta_3) \in E_q(\overline{\mathbb{Q}})$ such that

$$\Phi_{m_i}(\alpha_i, \beta_i) = 0 \text{ for } i = 1, 2, 3.$$

Remark 4.3.2. *We clarify that such an "avoiding curve" can be found also when C is contained in a proper weakly special subvariety of $Y(1)^3$, as we show below.*

The proof is, at least for most of the first part, similar to that of Theorem 4.2.1; here we recall the main arguments, focusing on the differences. The value M is chosen sufficiently big (namely, bigger than the degrees of the coordinates along C) in order to exploit the asymmetry of degrees together with Pazuki's Theorem 3.5.4.

Proof. For the first part of the proof, we need to make sure that our curves C and E_q are sufficiently "far from being weakly special" in order to apply the Pila-Wilkie Theorem 2.2.5.

Let us first deal with the case of C contained in a proper weakly special subvariety of $Y(1)^3$. There are two possibilities.

- If one coordinate, say x_1, is constant on the whole of C and equal to a value j_C, then E can be chosen as any curve with x_1 constant, assuming a value j_E such that the moduli j_C and j_E are not isogenous.
- If C instead has no constant coordinate but it is contained in a proper special subvariety, say satisfying identically

  $$\Phi_n(x_1, x_2) = 0$$

  for some positive integer n, then E can be chosen as a curve defined by the equations

  $$\begin{cases} x_1 = j_1 \\ x_2 = j_2 \end{cases}$$

  where j_1 and j_2 are two non isogenous moduli.

Remark 4.3.3. *In both the two cases above the curve E is definitely not "generic". It is possible to choose E in a more flexible family, adapting the proof below; we would need to suitably choose a restricted subset of coordinates for the application of the Ax-Schanuel Theorem for the j-function, in a fashion similar to that of the proof of Theorem 4.2.1. We do not pause on this.*

We can now assume C not contained in any proper weakly special subvariety of $Y(1)^3$. On the other hand, by the very same arguments of Remark 4.2.2, the curve E_q is not contained in a proper weakly special subvariety either: indeed, no coordinate is constant along E_q and the "asymmetry" between the degrees in the first, or second, and third coordinates implies that the only admissible special subvariety containing E_q is given by the zero locus of

$$\Phi_n(x_1, x_2) = 0$$

for some positive integer n. Whenever $q \neq 0$, then, by irreducibility of the modular curves, such zero locus cannot contain the line ℓ'

$$\{(t+q, t) \in Y(1)^2 \mid t \in Y(1)\}.$$

Hence, neither C nor E_q are contained in a proper weakly special subvariety. We observe that C and E_q are "related" (in the sense of section 7.1) only for finitely many values of q; indeed, the line $\ell' \subseteq Y(1)^2$ defined above is "identically isogenous" to the curve C', defined as the projection of C onto the first two coordinates, only for finitely many values of q.
The same argument holds also for the first and third coordinate together; for the second and third coordinate together, we just need to observe that an application of Pazuki's Theorem 3.5.4 as in Remark 4.2.2 implies that an identical isogeny between the curve

$$\{(t, t^M) \in Y(1)^2 \mid t \in Y(1)\}$$

and the projection of C onto the second and third factor cannot occur as long as M is bigger than the ratio of the degrees of the second and third coordinates on C. Indeed, $M = h(t^M)/h(t)$ needs to tend to

$$\deg_{x_3}(C)/\deg_{x_2}(C)$$

for $h(t)$ tending to infinity.

Summing up the reasoning above, for most choices of q then C and E_q are not contained in any proper weakly special subvariety of $Y(1)^3$ and, moreover, they are "unrelated", in the sense that there are no positive integers m, n and $1 \leq i < k \leq 3$ such that

$$\Phi_m(\alpha_i, \beta_i) = 0$$

$$\Phi_n(\alpha_k, \beta_k) = 0$$

for infinitely many points $(\alpha_1, \alpha_2, \alpha_3) \in C$ and $(\beta_1, \beta_2, \beta_3) \in E_q$.

These conditions are exactly those required for the application of the Pila-Wilkie Theorem 2.2.5 in the proof of Theorem 4.2.1; we notice that the constants implied are uniform in q. We describe explicitly the outcome.
For a point

$$(t+q, t, t^M) \in E_q(\overline{\mathbb{Q}})$$

its height is bounded above by

$$Mh(t) + h(q) + 1.$$

We suppose that for a point $(\alpha_1, \alpha_2, \alpha_3) \in C(\overline{\mathbb{Q}})$ we have

$$\Phi_{m_i}(\alpha_i, \beta_i) = 0 \text{ for } i = 1, 2, 3.$$

If we set $m = \max\{m_1, m_2, m_3\}$, then

$$m < c \cdot (Mh(t) + h(q) + 1)^6$$

for a constant c depending on C and M only.
We explicitly use Pazuki's Theorem 3.5.4 and we obtain the following inequalities:

$$|h(t+q) - h(\alpha_1)| \leq 12 \log m \leq c' + 72 \log h(t) + 72 \log h(q)$$

$$|h(t) - h(\alpha_2)| \leq 12 \log m \leq c' + 72 \log h(t) + 72 \log h(q)$$

$$\left|h(t^M) - h(\alpha_3)\right| \leq 12 \log m \leq c' + 72 \log h(t) + 72 \log h(q)$$

where c' is an absolute constant depending on C and M only.
We recall that, by the classical Néron result of [46], we obtain the inequality

$$\left|\deg_{x_j}(C)h(\alpha_i) - \deg_{x_i}(C)h(\alpha_j)\right| < c_1 \cdot \sqrt{1 + h(\alpha_i)} \text{ for every } 1 \leq i \neq j \leq 3$$

where c_1 is an absolute constant depending on C only. If we set the constant (which will be needed to be independent of M later)

$$D = \max\{\deg_{x_1}(C), \deg_{x_2}(C), \deg_{x_3}(C)\}$$

we can deduce the inequality

$$h(\alpha_i) < D(c_2 + h(\alpha_j)) \text{ for every } 1 \leq i \neq j \leq 3$$

for some constant c_2 depending on C only.
We use these algebraic manipulations in order to bound $h(q)$ in terms of $h(t)$:

$$h(q) < h(t) + h(t+q) + 2 \leq h(t) + 2 + h(\alpha_1) + c' + 72 \log h(t) + 72 \log h(q)$$

where we are using Proposition 2.6.7 for the basic properties of the Weil height. We can in turn bound $h(\alpha_1)$ in terms of $h(\alpha_2)$ and $h(\alpha_1)$ in terms of $h(t)$, namely

$$h(\alpha_1) < D(c_2 + h(\alpha_2))$$

and

$$h(\alpha_2) \leq h(t) + c' + 72 \log h(t) + 72 \log h(q)$$

and summarising, we obtain

$$h(q) < D'(1 + h(t) + \log h(q))$$

for a suitable constant D' depending on C and M only. This also implies that there is some constant D_1, depending on C and M only, such that

$$h(q) < D_1(1 + h(t)).$$

Finally, we compare $h(t^M)$ and $h(t)$ using the comparison above for $h(\alpha_2)$ and $h(\alpha_3)$:

$$h(t^M) \leq h(\alpha_3) + c' + 72 \log h(t) + 72 \log h(q) < h(\alpha_3) + e_1 \log h(t)$$

for a constant e_1 depending on C and M only;

$$h(\alpha_3) < D(c_2 + h(\alpha_2))$$

$$h(\alpha_2) \leq h(t) + c' + 72 \log h(t) + 72 \log h(q) < h(t) + e_1 \log h(t).$$

Thus, we conclude

$$Mh(t) = h(t^M) < Dh(t) + e_2(1 + \log h(t))$$

where e_2 is a constant depending on C and M only, while D is the constant defined above, which (crucially) depends on C only.
This gives an absolute (ineffective) upper bound on $h(t)$ provided $M > D$. In turn, we can use such upper bound in order to bound $h(q)$ from above and we obtain a contradiction for q big enough (in the sense that the existence of a point on C and a point on E_q which are isogenous is inadmissible).

□

We quote, without proof, the generalisation of the Proposition above to higher dimensions.

Let $X \subseteq Y(1)^r$ be a subvariety; there is a variety $Y \subseteq Y(1)^r$ of codimension $1 + \dim X$ which "avoids" X, in the sense that no point of $X(\overline{\mathbb{Q}})$ is isogenous to any point of $Y(\overline{\mathbb{Q}})$.
Explicitly, if X is $r-k-1$-dimensional and not contained in any proper weakly special subvariety of $Y(1)^r$, we can choose (possibly after a permutation of the coordinates) a k-dimensional subvariety $Y_{\underline{q}}$ defined as

$$\{(t_1+q_1, t_1+q_2, t_2+q_3, \ldots, t_k+q_{k+1}, t_1^M+q_{k+2}, \ldots, t_k^M+q_{2k+1}, t_1^{M^2}+q_{2k+2}, \ldots) \mid t_1, \ldots, t_k \in Y(1)\}$$

for an appropriate choice of a vector of integers $\underline{q} = (q_1, q_2, \ldots, q_r)$ or, alternatively, just in order to obtain a $Y_{\underline{q}}$ which is not a "product" variety,

$$\{(t_1+q_1, t_1+q_2, t_2+t_1+q_3, \ldots, t_k+t_{k-1}+q_{k+1}, t_1^M+q_{k+2}, \ldots, t_k^M+q_{2k+1}, \ldots) \mid t_1, \ldots, t_k \in Y(1)\}.$$

Remark 4.3.4. *An analogue of the result above seems hard to obtain (at least with our methods) in $\mathcal{A}_g$: exploiting "asymmetries" seems crucial in all of the proofs above and thus the "product structure" of $Y(1)^n$ is a necessary ingredient.*

We point out that we are not able to obtain an analogue of any of these results if C is allowed to vary among all the varieties of a given "algebro-geometric degree" and given degree of field definition (or even with a fixed field of definition): as we have seen in the section 4.1, we are unable to find a curve in $Y(1)^2$ which "avoids" all of $Y(1)^2(\mathbb{Q})$.

The situation seems even worse if we put additional restrictions on the "locus" containing E. An example here could be clearer than a definition.

Question. *Let $C \subseteq Y(1)^3$ be a curve over $\overline{\mathbb{Q}}$ and let $V \subseteq Y(1)^3$ be a surface "unrelated" to C (in the sense that no "isogenous image" of C meets V in a positive dimensional variety). Can we find a curve $E \subseteq V$ that "avoids" C (in the same sense as above)?*

This type of questions seems very out of reach, at least with our methods (which are based on the asymmetry of degrees, which is definitely not a "general" method).

Chapter 5

Tangencies in powers of the multiplicative group

In this section we prove a result concerning "double unlikely intersections" which generalises the result obtained by Marché-Maurin in [36].
The content of this chapter is in collaboration with Capuano and Ottolini.

Theorem 5.0.1 (B.-Capuano-Ottolini)**.** *Let n be a positive integer and let $C \subseteq \mathbb{G}_{\mathrm{m}}^n$ be a curve defined over $\overline{\mathbb{Q}}$. Suppose that C is not contained in a proper special subvariety; then there are only finitely many points $P \in C$ such that there is a one-codimensional special subvariety H satisfying:*

1. *P is in $C \cap H$;*
2. *the tangent space $T_P(C)$ of C at P is included in the tangent space $T_P(H)$ of H at P (namely, the curve C is tangent to H at P).*

Remark 5.0.2. *We point out that assuming C not contained in a proper special subvariety is a natural assumption; we see right below, in section 5.1, that discussing unlikely intersections restricted to proper special subvarieties of $\mathbb{G}_{\mathrm{m}}^n$ is the same as discussing unlikely intersections in $\mathbb{G}_{\mathrm{m}}^{n'}$ for some $n' < n$.*
Indeed, such an assumption is necessary: if C was contained in a proper special subvariety then it would also be contained in a one-codimensional proper special subvariety H and thus any point $P \in C \cap H$ would satisfy the two conditions above (i.e. C would be tangent to H at every point).

The Theorem by Marché-Maurin is the statement of Theorem 5.0.1 when $n = 2$; their proof generalises without any additional arguments whenever C is not contained in a proper weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^n$: in this case the height of any point in the intersection with one-codimensional special subvarieties has height bounded above by a result of Habegger (in [27]).
This height bound is not available anymore (indeed, it is just false) if C is contained in a proper weakly special subvariety and the proof is therefore quite different. We prove the height upper bound of Proposition 5.4.1 (depending on the "complexity" of the relevant one-codimensional

special subvariety) using Puiseux series expansions in order to bound the height at each individual place.

We point out that the statement of Theorem 5.0.1 becomes much easier if we require for a "triple intersection" rather than a "double" one, essentially because the exponential function exp satisfies a degree 1 differential equation.

We can also obtain a related "avoiding problem" exploiting the same arguments occurring in the proof of Theorem 5.0.1.
Let $\mathcal{F}$ be a "continuous family" of curves of $\mathbb{G}_{\mathrm{m}}^n$ (i.e. a subvariety of an appropriate Chow variety). We denote with $T(H)$ the set of curves $C \in \mathcal{F}$ such that C is tangent to a one-codimensional special subvariety H at some point.

Proposition 5.0.3. *Let us suppose that, for one-codimensional special subvariety H*

$$\textit{the set } T(H) \textit{ is not Zariski-dense in } \mathcal{F}.$$

Then, for "most" choices of $C \in \mathcal{F}$, there is no one-codimensional special subvariety H such that C is tangent to H at any point.

The hypotheses on the sets $T(H)$ are seen to be necessary: we can consider any family which is "identically tangent" to a single one-codimensional special subvariety H.

The proof is rather similar to that of Theorem 5.0.1; the main difficulty in extending our arguments is the following technical statement.

For any branch of any $C \in \mathcal{F}(\overline{\mathbb{Q}})$, the constants c and c_0 implied in its Puiseux series expansions satisfy

$$h(c), h(c_0) < \delta(1 + h(C))$$

for a constant δ depending on $\mathcal{F}$ only.

We refer to section 5.5 for the meaning of the constants c and c_0, which control the growth of the coefficients of the relevant Puiseux series expansion; with $h(C)$ we mean the height of the point representing C in $\mathcal{F}$.
We do not give proof of this "avoiding problem".

5.1 Special subvarieties and weakly special subvarieties

Let us describe the special subvarieties of the ambient variety $\mathbb{G}_{\mathrm{m}}^n$ for n a positive integer and $\mathbb{G}_{\mathrm{m}}$ the multiplicative group. The ground field is always assumed to be a subfield of $\overline{\mathbb{Q}}$. We fix coordinates $x_1, \ldots, x_n$, one for each factor of the product $\mathbb{G}_{\mathrm{m}}^n$. The special subvarieties in this context are exactly the irreducible algebraic subgroups translated by a torsion point, as

we have seen in section 3.1; they are precisely the irreducible components of the zero loci of systems of equations of the shape

$$\begin{cases} x_1^{p_{1,1}} \cdot x_2^{p_{2,1}} \cdot \ldots \cdot x_n^{p_{n,1}} = 1 \\ x_1^{p_{1,2}} \cdot x_2^{p_{2,2}} \cdot \ldots \cdot x_n^{p_{n,2}} = 1 \\ \ldots \\ x_1^{p_{1,m}} \cdot x_2^{p_{2,m}} \cdot \ldots \cdot x_n^{p_{n,m}} = 1 \end{cases}$$

for $p_{i,j}$, with $1 \leq i \leq n$ and $1 \leq j \leq m$, integers. We observe that the morphism

$$\begin{cases} x_1 \to x_1^{a_{1,1}} \cdot x_2^{a_{2,1}} \cdot \ldots \cdot x_n^{a_{n,1}} \\ x_2 \to x_1^{a_{1,2}} \cdot x_2^{a_{2,2}} \cdot \ldots \cdot x_n^{a_{n,2}} \\ \ldots \\ x_n \to x_1^{a_{1,n}} \cdot x_2^{a_{2,n}} \cdot \ldots \cdot x_n^{a_{n,n}} \end{cases}$$

is indeed a group automorphism whenever the $a_{i,j}$, with $1 \leq i, j \leq n$, are integers such that the $n \times n$ matrix with (i,j)-th entry $a_{i,j}$ is invertible (i.e. it has determinant equal to ± 1). Hence, using a suitable group automorphism (defined over $\mathbb{Q}$) and possibly permuting coordinates, a special subvariety will be an irreducible component of a system of the form

$$\begin{cases} x_1^{a_1} = 1 \\ x_2^{a_2} = 1 \\ \ldots \\ x_m^{a_m} = 1 \end{cases}$$

where the a_j, with $1 \leq j \leq m$, are integers. Equivalently, discarding the a_j which are equal to 0 and possibly decreasing m, any special subvariety H (under an automorphism described above and a permutation of coordinates) is of the shape

$$\begin{cases} x_1 = \omega_1 \\ x_2 = \omega_2 \\ \ldots \\ x_m = \omega_m \end{cases}$$

for roots of unity ω_j for every $1 \leq j \leq m$. Notice that $\mathbb{Q}(\omega_1, \omega_2, \ldots, \omega_m)$ is then the field of definition of H. This argument shows that any special subvariety H of codimension m is isomorphic to some $\mathbb{G}_{\mathrm{m}}^{n-m}$ by an isomorphism (defined over the field of definition of H) that preserves special subvarieties; one defines the special subvarieties of H as the intersections of H and any special subvariety of $\mathbb{G}_{\mathrm{m}}^n$.

Special points (i.e. zero-dimensional special subvarieties) are precisely the points $(\omega_1, \ldots, \omega_n) \in \mathbb{G}_{\mathrm{m}}^n$ with the ω_j, for $1 \leq j \leq n$, roots of unity. One-codimensional special subvarieties are precisely the zero loci of a single equation of the shape

$$x_1^{p_1} \cdot x_2^{p_2} \cdot \ldots \cdot \ldots \cdot x_n^{p_n} = \omega$$

where the p_i, for $1 \leq i \leq n$, are coprime integers (in particular, not all equal to zero), and ω is a root of unity.

We now shift our attention to weakly special subvarieties of $\mathbb{G}_{\mathrm{m}}^n$, even though our statement only involves special subvarieties. We have seen already during the proof of Theorem 3.6.7 that the presence of positive dimensional weakly special subvarieties is a major obstacle and here they complicate the situation as well.

The weakly special subvarieties of $\mathbb{G}_{\mathrm{m}}^n$ are exactly the irreducible algebraic subgroups translated by a point (which might not be a torsion point) and they are precisely the irreducible components of the zero locus of systems of equations of the shape

$$\begin{cases} x_1^{p_{1,1}} \cdot x_2^{p_{2,1}} \cdot \ldots \cdot x_n^{p_{n,1}} = c_1'' \\ x_1^{p_{1,2}} \cdot x_2^{p_{2,2}} \cdot \ldots \cdot x_n^{p_{n,2}} = c_2'' \\ \ldots \\ x_1^{p_{1,m}} \cdot x_2^{p_{2,m}} \cdot \ldots \cdot x_n^{p_{n,m}} = c_m'' \end{cases}$$

for $p_{i,j}$, with $1 \leq i \leq n$ and $1 \leq j \leq m$, integers and c_j'', for $1 \leq j \leq m$, arbitrary nonzero scalars. Again, using a suitable group automorphism (which is defined over $\mathbb{Q}$) and possibly permuting coordinates, as we described above for special subvarieties, any weakly special subvariety will be an irreducible component of a system of the form

$$\begin{cases} x_1^{a_1} = c_1' \\ x_2^{a_2} = c_2' \\ \ldots \\ x_m^{a_m} = c_m' \end{cases}$$

where the a_j, with $1 \leq j \leq m$, are integers and the c_j', for $1 \leq j \leq m$, are arbitrary nonzero scalars. Equivalently, discarding the a_j which are equal to 0 and possibly decreasing m, any weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^n$ (under an automorphism described above and a permutation of coordinates) is of the shape

$$\begin{cases} x_1 = c_1 \\ x_2 = c_2 \\ \ldots \\ x_m = c_m \end{cases}$$

for nonzero scalars c_j, for every $1 \leq j \leq m$. Again, we notice that $\mathbb{Q}(c_1, c_2, \ldots, c_m)$ is then the field of definition of our weakly special subvariety. This shows that any weakly special subvariety of codimension m is isomorphic (over its field of definition) to some $\mathbb{G}_{\mathrm{m}}^{n-m}$; special subvarieties are not preserved by such an isomorphism and moreover it is not even clear what should be the notion of a special subvariety of a weakly special subvariety.

Weakly special points (i.e. zero-dimensional weakly special subvarieties) are just all the points of $\mathbb{G}_{\mathrm{m}}^n$. One-codimensional weakly special subvarieties are precisely the zero loci of a single equation of the shape

$$x_1^{p_1} \cdot x_2^{p_2} \cdot \ldots \cdot \ldots \cdot x_n^{p_n} = c$$

where the p_i, for $1 \leq i \leq n$, are coprime integers (in particular, not all equal to zero) and c is a nonzero scalar.

5.2 Bounds on the degree

In this section we prove Theorem 5.0.1 assuming that C is not contained in any proper weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^n$; most of the observations that we make will also be useful in the general case.

We start by bounding the degree over $\mathbb{Q}$ of the field of definition of a point $P \in C \cap H$, where H is a one-codimensional special subvariety and C is tangent to P at H; we can even relax the condition on H and just ask for it to be weakly special: given a one-codimensional weakly special subvariety W, the degree of the field of definition of a point $P \in C \cap W$ such that C is tangent to W at P is bounded independently of W.
We remark that, for these first proofs, it will be crucial for C not to be contained in any proper weakly special subvariety and we assume so for now.

Proposition 5.2.1. *Suppose that the curve $C \subseteq \mathbb{G}_{\mathrm{m}}^n$ is not contained in any proper weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^n$ and that C is defined over the number field K. There is a constant $\delta_1 > 0$, depending only on C, such that, for every one-codimensional weakly special subvariety $W \subseteq \mathbb{G}_{\mathrm{m}}^n$ and every point $P \in C \cap W$ with C tangent to W at P:*

$$[K(P) : K] \leq \delta_1.$$

Let $x_1, \ldots, x_n$ denote the coordinates of $\mathbb{G}_{\mathrm{m}}^n$ (one coordinate for each factor, as usual); we can restrict these coordinate functions to C and we also obtain corresponding differentials $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$. We can thus construct the morphism

$$f : C \to \mathbb{P}_{n-1};$$

$$f(x_1, \ldots, x_n) = \left(\frac{\mathrm{d}x_1}{x_1}, \ldots, \frac{\mathrm{d}x_n}{x_n} \right).$$

Actually, f might not be defined at a point P as long as all the differentials $\mathrm{d}x_1, \ldots, \mathrm{d}x_n$ vanish; we get rid of this issue by removing all the singular points from C (in particular, we are not asking for C to be closed in $\mathbb{G}_{\mathrm{m}}^n$). The individual $\mathrm{d}x_i/x_i$ are not well-defined values, but the ratios of two such values are (since the ratio of two differentials on C is just a rational function on C). To see how to make sense of f at a smooth point $P \in C$, we can pick any nonzero element t of the tangent space $T_P(C)$; then for some $1 \leq i \leq n$ we have that $(\mathrm{d}x_i)(t)$ is nonzero (since P is a smooth point) and then

$$f(P) = \left(\frac{(\mathrm{d}x_1)(t)}{x_1(P)}, \ldots, \frac{(\mathrm{d}x_n)(t)}{x_n(P)} \right).$$

This value is independent of t as all the nonzero elements of $T_P(C)$ are just multiples of t (again, by smoothness of P). Notice that f is just the logarithmic derivative and therefore we expect multiplicative relations to translate into linear relations via f; indeed the following holds.

Proposition 5.2.2. *Let P be a point of $C \cap W$, where W is the one-codimensional weakly special subvariety defined by $x_1^{p_1} \dots x_n^{p_n} = c$ for p_i coprime integers and c a nonzero scalar. Assume that C is tangent to W at P. Then*

$$p_1 \frac{dx_1}{x_1} + p_2 \frac{dx_2}{x_2} + \dots + p_n \frac{dx_n}{x_n} = 0$$

at P. Equivalently, $f(P)$ is contained in the hyperplane $L \subseteq \mathbb{P}_{n-1}$ given by the equation

$$p_1 z_1 + \dots + p_n z_n = 0$$

where $z_1, \dots, z_n$ are the (usual) coordinates of $\mathbb{P}_{n-1}$.

Proof. Since $T_P(C)$ is contained in $T_P(W)$ this implies that the differential (on W)

$$\mathrm{d}(x_1^{p_1} \dots x_n^{p_n} - c)$$

which vanishes on the whole of $T_P(W)$, must vanish (as a differential on C) on the whole of $T_P(C)$, hence it is zero at P. Expanding using the Leibniz rule we obtain

$$\mathrm{d}(x_1^{p_1} \dots x_n^{p_n} - c) = x_1^{p_1} \dots x_n^{p_n} \left(p_1 \frac{\mathrm{d}x_1}{x_1} + p_2 \frac{\mathrm{d}x_2}{x_2} + \dots + p_n \frac{\mathrm{d}x_n}{x_n} \right)$$

and then we are done, as the x_i are nonzero. □

Now, let $P \in C \cap W$ be a point such that C is tangent to W at P. Then $f(P)$ is in $L \cap f(C)$ and the latter is a finite set, unless $f(C)$ is contained in L; we suppose that this is not the case. We have two possibilities:

1. $f(C)$ is a point: this means that $f(C)$ is exactly $\{f(P)\}$ and therefore $f(C)$ is contained in L;

2. $f(C)$ is a curve: observe that in this case there is an upper bound on the cardinality $|L \cap f(C)|$ which does not depend on L; such an upper bound is the usual "algebro-geometric" degree d of $f(C)$ with respect to the line bundle $O(1)$ (and the following holds: any hyperplane intersects $f(C)$ in at most d points and a generic hyperplane intersects $f(C)$ in exactly d points).
   As L is defined over $\mathbb{Q}$ and $f(C)$ is defined over K, this means that

   $$[K(f(P)) : K] \leq |L \cap f(C)| \leq d$$

   and hence $[K(P) : K] \leq d \cdot \deg f$, where $\deg f$ is the degree of the morphism of curves given by $f : C \to f(C)$.

We are left with the exceptional case of $f(C)$ contained in L. In this case we observe that the differential

$$\mathrm{d}(x_1^{p_1} x_2^{p_2} \dots x_n^{p_n}) = p_1 \frac{\mathrm{d}x_1}{x_1} + p_2 \frac{\mathrm{d}x_2}{x_2} + \dots + p_n \frac{\mathrm{d}x_n}{x_n}$$

is constantly zero on C and this means that $x_1^{p_1}x_2^{p_2}\dots x_n^{p_n}$ is constant. Hence C is contained in a weakly special subvariety given by the equation

$$x_1^{p_1}x_2^{p_2}\dots x_n^{p_n} = c'.$$

For some $c' \in \mathbb{G}_{\mathrm{m}}$. We will deal later with this case, that we have excluded for now. This concludes the proof of Proposition 5.2.1.

We are now ready to prove Theorem 5.0.1 in the case of C not contained in a proper weakly special subvariety. Let H be a one-codimensional special subvariety and let $P \in C \cap H$. A Theorem of Bombieri-Masser-Zannier [10] (which was then generalised by Habegger in [27]) tells us the following.

The height $h(P)$ is bounded by a constant independent of H.

This, in particular, proves, as long as C is not contained in a proper weakly special subvariety, finiteness: we have bounded degree and bounded height for P, hence only finitely many possibilities by Northcott's Theorem.

5.3 General reductions for weakly special subvarieties

We now suppose that C is contained in a proper weakly special subvariety W and moreover we suppose that W has minimal dimension; we can also assume that C is not contained in any proper special subvariety.
Let K be the number field over which C is defined. We argue that W can be taken defined over K. We consider the intersection of all the conjugates of W via elements of $\mathrm{Gal}\,(\overline{\mathbb{Q}}/K)$ and observe that the irreducible components of arbitrary intersections of weakly special subvarieties are weakly special subvarieties (indeed, by Hilbert's Basis Theorem, any infinite system of equations is equivalent to a finite subsystem). Let W' be an irreducible component of such intersection containing C: we have that $W' \subseteq W$ is a weakly special subvariety, but then $W' = W$ by minimality and therefore W is equal to all of its conjugates via $\mathrm{Gal}\,(\overline{\mathbb{Q}}/K)$ (i.e. it is defined over K).

Considering an appropriate automorphism (defined over $\mathbb{Q}$) of $\mathbb{G}_{\mathrm{m}}^n$ together with a permutation of the coordinates, as described above, we can posit W to be an irreducible component of a system of the shape

$$\begin{cases} x_1^{a_1} = c_1' \\ x_2^{a_2} = c_2' \\ \dots \\ x_m^{a_m} = c_m' \end{cases}$$

with, for $1 \le j \le m$, the a_j integers and the c_j' elements of K^*. Therefore, possibly permuting the coordinates and ignoring the a_j equal to zero, the irreducible variety W will be given by

the system of equations

$$\begin{cases} x_1 = c_1 \\ x_2 = c_2 \\ \dots \\ x_m = c_m \end{cases}$$

with c_j, for $1 \leq j \leq m$, nonzero elements of $\overline{\mathbb{Q}}$. Actually, the c_j all lie in K, since W is defined over such a field.

We assume from now on that W is given by such equations, $C \subseteq W$ and both are defined over K. We use the assumption of C not being contained in a proper special subvariety of $\mathbb{G}_{\mathrm{m}}^n$.

Lemma 5.3.1. *The values $c_1, c_2, \dots, c_m$ are multiplicatively independent.*

Proof. Suppose that for integers $q_1, q_2, \dots, q_m$, not all zero, we have

$$c_1^{q_1} c_2^{q_2} \dots c_m^{q_m} = 1$$

and then W is contained in the zero locus of

$$x_1^{q_1} x_2^{q_2} \dots x_m^{q_m} = 1$$

which is a finite union of proper special subvarieties; since $C \subseteq W$ we obtain a contradiction. □

We now wish to restrict our attention to some $\mathbb{G}_{\mathrm{m}}^{n-m}$ (with coordinates $y_1, \dots, y_{n-m}$) and we consider the following injection:

$$\iota : \mathbb{G}_{\mathrm{m}}^{n-m} \to \mathbb{G}_{\mathrm{m}}^n$$

$$\iota(y_1, \dots, y_{n-m}) = (c_1, \dots, c_m, y_1, \dots, y_{n-m})$$

and notice that $\iota(\mathbb{G}_{\mathrm{m}}^{n-m})$ is exactly W, so that $\iota^{-1}(C)$ is a curve.

We use again minimality of W.

Lemma 5.3.2. *$\iota^{-1}(C)$ is not contained in any proper weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^{n-m}$.*

Proof. Suppose $\iota^{-1}(C) \subseteq W'$ for W' a proper weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^{n-m}$ i.e. the translate of an algebraic subgroup; then the image $\iota(W')$ is again the translate of an algebraic subgroup and it is strictly contained in W, but this contradicts minimality of the dimension of W. □

We now characterise the preimages $\iota^{-1}(H)$ whenever H is a one-codimensional special subvariety of $\mathbb{G}_{\mathrm{m}}^n$. Such a special subvariety is the zero locus of an equation of the shape

$$x_1^{a_1} x_2^{a_2} \dots x_n^{a_n} = \omega$$

for a_j, with $1 \leq j \leq n$, coprime integers and ω a root of unity. The equation for $\iota^{-1}(H)$ is then

$$c_1^{a_1} c_2^{a_2} \dots c_m^{a_m} \cdot y_1^{a_{m+1}} y_2^{a_{m+2}} \dots y_{n-m}^{a_n} = \omega.$$

We notice that $\iota^{-1}(H)$ might not be irreducible (this happens precisely when the integers $a_{m+1}, a_{m+2}, \dots, a_n$ are not coprime). We have two possibilities:

1. all the a_j, for $m+1 \leq j \leq n$, are zero; but then

   $$c_1^{a_1} c_2^{a_2} \dots c_m^{a_m} = \omega$$

   and hence $c_1, c_2, \dots, c_m$ are multiplicatively dependent by the Lemma above (since not all of the $a_1, a_2, \dots, a_m$ are zero), which we know is not the case, unless $\iota^{-1}(H)$ is empty;

2. some a_j, with $m+1 \leq j \leq n$, is nonzero; then $\iota^{-1}(H)$ is a finite union of one-codimensional weakly special subvarieties of $\mathbb{G}_{\mathrm{m}}^{n-m}$.

This characterisation leads us to give the following definition. We think of n' as of $n-m$.

Definition. *Let m be a natural number and let $\underline{c}$ be an m-tuple of multiplicatively independent nonzero scalars $(c_1, c_2, \dots, c_m)$. A $\underline{c}$-special hypersurface of $\mathbb{G}_{\mathrm{m}}^{n'}$ is an irreducible component of the zero locus of an equation of the shape*

$$c_1^{q_1} c_2^{q_2} \dots c_m^{q_m} \cdot x_1^{p_1} x_2^{p_2} \dots x_{n'}^{p_{n'}} = \omega$$

for integers p_i, with $1 \leq i \leq n'$, integers q_j, with $1 \leq j \leq m$, such that some p_i is nonzero and such that

$$gcd(p_1, p_2, \dots, p_{n'}, q_1, q_2, \dots, q_m) = 1$$

and ω a root of unity.

Notice that $\underline{c}$-special hypersurfaces are one-codimensional weakly special subvarieties. This notion allows us to focus our attention on unlikely intersections in $\mathbb{G}_{\mathrm{m}}^{n-m}$ and forget about $\mathbb{G}_{\mathrm{m}}^{n}$, thanks to the following statement. We refer to the inclusion $\iota : \mathbb{G}_{\mathrm{m}}^{n-m} \to \mathbb{G}_{\mathrm{m}}^{n}$ described above.

Lemma 5.3.3. *Let H be a one-codimensional special subvariety of $\mathbb{G}_{\mathrm{m}}^{n}$ and let $P \in C \cap H$ such that C is tangent to H at P. Then $\iota^{-1}(C)$ is tangent to U at $\iota^{-1}(P)$, where U is a $\underline{c}$-special hypersurface of $\mathbb{G}_{\mathrm{m}}^{n-m}$.*

Proof. We prove that $T_P(W \cap H) = T_P(W) \cap T_P(H)$. Notice that W and H are translates of algebraic subgroups; if we denote with W_0 and H_0 the images of W and H by the translation that maps P into the origin, the assertion above amounts to saying that the Lie algebra of (the irreducible component of) $W_0 \cap H_0$ (containing the origin) coincides with the intersection of the Lie algebras of W_0 and H_0. This is true for the complex Lie group $\mathbb{G}_{\mathrm{m}}^{n}$: its Lie algebra can be identified with some $\mathbb{C}^n$ and its closed complex subgroups (which are all algebraic) correspond to rational subspaces of $\mathbb{C}^n$, which are closed by intersection.
We now have that $T_P(C) \subseteq T_P(H)$ by hypothesis and, since $C \subseteq W$, then $T_P(C) \subseteq T_P(W \cap H)$ by the argument above. Observe that ι^{-1} is an isomorphism between W and $\mathbb{G}_{\mathrm{m}}^{n-m}$, therefore $T_{\iota^{-1}(P)}(\iota^{-1}(C)) \subseteq T_{\iota^{-1}(P)}(\iota^{-1}(H))$ and take U to be the unique irreducible component of $\iota^{-1}(H)$ containing P (notice that $\iota^{-1}(H)$ is non-empty as $P \in W \cap H$). □

Thanks to the last two Lemmas, we see that Theorem 5.0.1 will now be a consequence of the following (we replace $\mathbb{G}_{\mathrm{m}}^{n-m}$ with $\mathbb{G}_{\mathrm{m}}^{n}$ for simplicity):

Theorem 5.3.4. *Let m and n be positive integers, let $\underline{c} = (c_1, c_2, \ldots, c_m)$ be a vector of multiplicatively independent nonzero elements of $\overline{\mathbb{Q}}$ and let $C \subseteq \mathbb{G}_{\mathrm{m}}^{n}$ be a curve defined over $\overline{\mathbb{Q}}$ which is not contained in any proper weakly special subvariety. Then there are only finitely many points $P \in C$ such that there is a $\underline{c}$-special hypersurface U such that:*

1. *P is in $C \cap U$;*
2. *the tangent space $T_P(C)$ of C at P is included in the tangent space $T_P(U)$ of U at P (namely, the curve C is tangent to U at P).*

5.4 Bounds on the height: functoriality

In this section we make extensive use of the Weil height. We always assume our varieties to be embedded in a projective space of which they inherit the usual projective Weil height. For instance, we embed $\mathbb{G}_{\mathrm{m}}^{n}$ into $\mathbb{P}_n$ via

$$(x_1, \ldots, x_n) \to (x_1, \ldots, x_n, 1).$$

We prove the following statement.

Proposition 5.4.1. *Let n be a positive integer and let $C \subseteq \mathbb{G}_{\mathrm{m}}^{n}$ be a curve defined over the number field K; suppose that C is not contained in any proper weakly special subvariety of $\mathbb{G}_{\mathrm{m}}^{n}$. There is a constant $\delta_2 > 0$, depending only on C, with the following property. Suppose that C is tangent, at P, to a one-codimensional weakly special subvariety W contained in the zero locus of*

$$x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n} = c$$

where the q_i, for $1 \leq i \leq n$, are integers, not all zero, and c is a nonzero element of $\overline{\mathbb{Q}}$. Then:

$$h(P) \leq \delta_2(1 + h(p_1) + h(p_2) + \ldots + h(p_n)).$$

We consider, as above, the morphism

$$f : C \to \mathbb{P}_{n-1}$$

$$f(x_1, \ldots, x_n) = \left(\frac{\mathrm{d}x_1}{x_1}, \ldots, \frac{\mathrm{d}x_n}{x_n}\right)$$

possibly removing the finitely many singular points from C and we recall that, by Proposition 5.2.2, $f(P)$ is contained in the hyperplane L given by the equation

$$p_1 z_1 + p_2 z_2 + \ldots + p_n z_n = 0$$

where $z_1, z_2, \ldots, z_n$ are the coordinates of $\mathbb{P}_{n-1}$. As before, if $f(C)$ was contained in L, this would imply that

$$\mathrm{d}(x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n}) = p_1 \frac{\mathrm{d}x_1}{x_1} + p_2 \frac{\mathrm{d}x_2}{x_2} + \ldots + p_n \frac{\mathrm{d}x_n}{x_n} = 0$$

on all of C and therefore C would be contained in a one-dimensional weakly special subvariety. Thus P is contained in the finite set $f^{-1}(f(C) \cap L)$ (as before, f is nonconstant since $f(C)$ is not contained in L).

It is now classical (as a consequence of 2. of Theorem 2.7.2, for instance) that, for a finite map of curves $g : C_1 \to C_2$, once we specify the projective embeddings, there is a constant $\eta > 0$ such that

$$h(Q_1) \leq \eta(1 + h(Q_2))$$

for all $Q_1 \in C_1(\overline{\mathbb{Q}})$ and $Q_2 \in C_2(\overline{\mathbb{Q}})$ such that $g(Q_1) = Q_2$. Therefore, we just need to bound the height of any element of $f(C) \cap L$ (notice that C and $f(C)$ are already embedded in $\mathbb{P}_n$ and $\mathbb{P}_{n-1}$ respectively). This amounts to the following result.

Lemma 5.4.2. *Let s be a positive integer and $C' \subseteq \mathbb{P}_s$ be a curve defined over $\overline{\mathbb{Q}}$. There is a constant $\delta' > 0$ with the following property. If $L_{\underline{p}}$ is a hyperplane given by the equation*

$$p_0 z_0 + p_1 z_1 + \ldots + p_s z_s = 0$$

for $\underline{p} = (p_0, p_1, \ldots, p_s)$ with $p_0, p_1 \ldots, p_s$ elements of $\overline{\mathbb{Q}}$ and $z_0, z_1, \ldots, z_s$ the coordinates of $\mathbb{P}_s$ and if $L_{\underline{p}}$ does not contain C', then, for any point $P \in C' \cap L$

$$h(P) \leq \delta'(1 + h(p_0) + h(p_1) + \ldots + h(p_s)).$$

Proof. We consider the Grassmanian G which parameterises hyperplanes in $\mathbb{P}_s$, which is isomorphic to $\mathbb{P}_s$ (that we can think of as with coordinates $p_0, p_1, \ldots, p_s$). Let $V \subseteq G$ be the subset consisting of $\underline{p} \in G$ such that $L_{\underline{p}}$ does not contain C; then V is an open subset of $\mathbb{P}_s$ (since the requirement of containing a prescribed point is a closed condition and therefore $G \setminus V$ is an intersection of closed sets).

Our strategy consists in constructing some morphisms which behave well with respect to heights in order to bound the height of P from the height of $\underline{p}$.

We can suppose that C' is closed in $\mathbb{P}_s$ (this would just enlarge $C' \cap L_{\underline{p}}$). By Bézout's Theorem, there is a positive integer d such that, for any $\underline{p} \in V$, we have that $L_{\underline{p}}$ intersects C' in precisely d points (counted with multiplicity). This gives rise to a morphism

$$\varphi_1 : V \to (\mathbb{P}_s)^d / S_d$$

$$\varphi_1(\underline{p}) = C' \cap_m L_{\underline{p}}$$

where $(\mathbb{P}_s)^d / S_d$ is the d-th symmetric power of $\mathbb{P}_s$ (and it is a variety) and $C' \cap_m L_{\underline{p}}$ denotes the intersection of C' and $L_{\underline{p}}$ taking into account multiplicities. Since $(\mathbb{P}_s)^d / S_d$ is the quotient of $(\mathbb{P}_s)^d$ by the finite group S_d, the quotient map defines a finite morphism

$$\varphi_2 : (\mathbb{P}_s)^d \to (\mathbb{P}_s)^d / S_d$$

$$\varphi_2(P_1, \ldots, P_d) = (P_1, \ldots, P_d)_{no}$$

in the sense that the d-uple is ordered in the LHS and non-ordered in the RHS (we use the subscript $_{no}$ to denote this). Finally, we consider the projection on the first coordinate

$$\varphi_3 : (\mathbb{P}_s)^d \to \mathbb{P}_s;$$

$$\varphi_3(P_1, \ldots, P_d) = P_1.$$

Let now be $P \in C' \cap L_{\underline{p}}$. We can choose points $Q_1, \ldots, Q_{d-1}$ in $\mathbb{P}_s$ such that $(P, Q_1, \ldots, Q_{d-1})_{no} = C' \cap_m L_{\underline{p}}$. We have

$$\varphi_1(\underline{p}) = (P, Q_1, \ldots, Q_{d-1})_{no}$$

$$\varphi_2(P, Q_1, \ldots, Q_{d-1}) = (P, Q_1, \ldots, Q_{d-1})_{no}$$

$$\varphi_3(P, Q_1, \ldots, Q_{d-1}) = P.$$

It is classical (and a consequence of 1. of Theorem 2.7.2) that if $\varphi : V_1 \to V_2$ is a morphism of varieties (embedded in a projective space from which they inherit height functions) then there is some $\rho > 0$ such that

$$h(Q_2) < \rho(1 + h(Q_1))$$

whenever $\varphi(Q_1) = Q_2$. Moreover, if $\varphi : V_1 \to V_2$ is finite, then there is some $\theta > 0$ such that

$$h(Q_1) < \theta(1 + h(Q_2))$$

whenever $\varphi(Q_1) = Q_2$ and $\varphi^{-1}(Q_2)$ is a finite set. It is classical that $(\mathbb{P}_s)^d$ and $(\mathbb{P}_s)^d/S_d$ can be embedded in some projective space (and V is already embedded in G). We fix some embedding from which we obtain height functions; by the argument outline before, there are constants η_1, η_2, η_3 for $\varphi_1, \varphi_2, \varphi_3$ respectively such that, for any choice of $\underline{p}, P, Q_1, \ldots, Q_{d-1}$ as above:

$$h((P, Q_1, \ldots, Q_{d-1})_{no}) < \eta_1(1 + h(\underline{p}))$$

$$h(P, Q_1, \ldots, Q_{d-1}) < \eta_2(1 + h((P, Q_1, \ldots, Q_{d-1})_{no}))$$

$$h(P) < \eta_3(1 + h(P, Q_1, \ldots, Q_{d-1}))$$

and since

$$h(\underline{p}) = h(p_1) + \ldots + h(p_s)$$

we are done. □

5.5 Bounds on the height: Puiseux series

In this section we give an alternative proof of Lemma 5.4.2 in the case of the hyperplane being defined over $\mathbb{Q}$, using Puiseux series.
We rephrase the Lemma in this new version here.

Lemma 5.5.1. *Let s be a positive integer and $C' \subseteq \mathbb{P}_s$ be a curve defined over $\overline{\mathbb{Q}}$. There is a constant $\delta' > 0$ with the following property. If $L_{\underline{p}}$ is an hyperplane given by the equation*

$$p_0 z_0 + p_1 z_1 + \ldots + p_s z_s = 0$$

for $\underline{p} = (p_0, p_1, \ldots, p_s)$ with $p_0, p_1, \ldots, p_s$ elements of $\mathbb{Q}$ and $z_0, z_1, \ldots, z_s$ the coordinates of $\mathbb{P}_s$ and $L_{\underline{p}}$ does not contain C', then, for any point $P \in C' \cap L$

$$h(P) \leq \delta'(1 + h(p_0) + h(p_1) + \ldots + h(p_s)).$$

We start by recalling some facts about Puiseux series of algebraic functions (see the paper by Van der Poorten [57]). We point out that, in order to estimate the height, we prefer to expand "around infinity" rather than in a neighbourhood of zero (i.e. we take series in $\mathbb{C}((t^{-1}))$ rather than in $\mathbb{C}((t))$).

Let K be a number field and let $f(x, y) \in K[x, y]$ be an irreducible polynomial of degree d in y, i.e. as a polynomial in $K(x)[y]$. Then there are exactly d elements $y_1, y_2, \ldots, y_d \in \overline{K(x)}$ such that $f(x, y_i) = 0$ for any $1 \leq i \leq d$.

For an appropriate positive integer Q, we can write any y_i, for any $1 \leq i \leq d$, as

$$y_i = a_r x^{r/Q} + a_{r-1} x^{(r-1)/Q} + a_{r-2} x^{(r-2)/Q} + \ldots \text{ an element of } \mathbb{C}((x^{-1/Q}))$$

for an infinite sequence $a_r, a_{r-1}, a_{r-2}, \ldots$ of complex numbers and, since $f(x, y_i) = 0$, we obtain remarkably, by a consideration of Eisenstein, that there are a number field K' and two algebraic integers $\alpha_0, \alpha \in K'$ such that

$$\alpha_0 \cdot \alpha^i \cdot a_{r-i}$$

is an algebraic integer of K' for every $i \geq 0$ (this is at the beginning of the third chapter of the expository article by Van der Poorten [57]). Moreover, the Puiseux series expansion above converges (in $\mathbb{C}$ - this is the so-called Newton-Puiseux Theorem) whenever $|x|$ is big enough, after choosing a suitable Q-th root of x. We refer to Nowak [58] for this.

Let us see how these considerations fit in our picture. Let us fix an open affine $\mathbb{A}^s \subseteq \mathbb{P}_s$ defined for instance by $z_0 \neq 0$, where $z_0, z_1, \ldots, z_s$ are the coordinates of $\mathbb{P}_s$, and let us set $y_i = z_i/z_0$ for every $1 \leq i \leq s$ the coordinates on such $\mathbb{A}^s$ when restricted to $\mathbb{A}^s \cap C'$. We let K be the number field over which C' is defined.
We write y_i, for any $2 \leq i \leq s$, as a Puiseux series of y_1; this is possible since y_i and y_1 are algebraically dependent since they are restricted to C' (and we might assume, by permuting coordinates, that y_1 is nonconstant on C'). We obtain a Puiseux series expansion

$$y_i = a_r y_1^{r/Q} + a_{r-1} y_1^{(r-1)/Q} + a_{r-2} y_1^{(r-2)/Q} + \ldots$$

for some positive integer Q, an integer r and coefficients $a_r, a_{r-1}, a_{r-2}, \ldots$ depending on i; moreover, we notice that y_i will admit d Puiseux series expansions, where d is the degree of y_i

over $K(y_1)$.
Each y_i (for any $1 \leq i \leq s$) admits Puiseux series expansions in y_1; while not strictly necessary, we write $y_1 = x^{Q'}$ for a suitable positive integer Q and we can then suppose that all the Puiseux series constructed as above are of the form

$$y_i = a_r x^r + a_{r-1} x^{r-1} + a_{r-2} x^{r-2} + \dots$$

for a suitable positive integer r and coefficients $a_r, a_{r-1}, a_{r-2}, \dots$ depending on i and the choice of the branch of y_i.

Let us track the dependence of the coefficients on i, namely fixing Puiseux series expansions for each y_i

$$y_i = a_{r,i} x^r + a_{r-1,i} x^{r-1} + a_{r-2,i} x^{r-2} + \dots .$$

For every i, as long as $|x|$ is big enough, the series above converges to the correct value, in this sense: if, as functions, x and y_i satisfy $f(x, y_i) = 0$ for a polynomial $f(t_1, t_2) \in \mathbb{C}[t_1, t_2]$, then, as the numbers above, $f(x, y_i) = 0$.
This argument indeed implies the following observations related to convergence.

- *Convergence over $\mathbb{C}$:* we consider a point $(y_1, y_2, \dots, y_s) \in C' \cap \mathbb{A}^s(\mathbb{C})$ such that $|y_1|$ is big enough. Then there is a choice of a complex value x (which is a Q'-th root of y_1) and suitable determinations of the Puiseux series such that each y_i (for $1 \leq i \leq s$) is the value of a convergent Puiseux series as above.
  Hence, all the points with a very big "infinite part" of the height of y_1 can be expressed, in each coordinate, by the Puiseux series described above.

- *Convergence over $\mathbb{C}_p$:* we now argue similarly, but for finite places. Indeed, by the observation by Eisenstein quoted by Van der Poorten, all the Puiseux series constructed above for the various y_i (with $1 \leq i \leq s$) have their coefficients in a common number field K' and the existence of nonzero numbers $\alpha_0, \alpha \in K'$ such that $\alpha_0 \cdot \alpha^j \cdot a_{r-j,i}$ are algebraic integers for every natural number j and every choice of $1 \leq i \leq s$ (and every determination for each y_i as well).
  Let us take an algebraic $x \in \overline{\mathbb{Q}}$ and let $\mathfrak{p}$ be a prime of a number field containing $K'(x)$. Then, if the valuation $v_{\mathfrak{p}}(x)$ is sufficiently negative, the power series above will converge p-adically: this can be even seen explicitly, as

  $$v_{\mathfrak{p}}(x) < -v_{\mathfrak{p}}(\alpha)$$

  suffices.
  Fixing an appropriate embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}_p$, where p is the rational prime number underlying $\mathfrak{p}$, we obtain a value y_i, that, as in the complex case, satisfies $f(x, y_i)$ for any algebraic relation satisfied identically (i.e. for x and y_i as function) over $\mathbb{C}_p$.
  We can therefore obtain the following: let $(y_1, y_2, \dots, y_s) \in C' \cap \mathbb{A}^s(\overline{\mathbb{Q}})$ and let x be a Q'-th root of y_1. Let K'' be the number field $K'(x, y_1, y_2, \dots, y_s)$ and let $\mathfrak{p}$ be a prime

of K'' lying above the rational prime p. We fix an inclusion $\overline{\mathbb{Q}} \subseteq \mathbb{C}_p$ so that $\mathfrak{p} \subseteq \overline{\mathbb{Q}}$ is exactly the preimage of the elements of $\mathbb{C}_p$ with positive valuation.
If $v_{\mathfrak{p}}(x)$ is sufficiently negative, the arguments above imply that there are suitable determinations of the Puiseux series such that each y_i (for $1 \leq i \leq s$) is the value of a convergent Puiseux series in $\mathbb{C}_p$; we are tacitly using that distinct Puiseux series coincide at only finitely many points (and they will not coincide whenever $v_{\mathfrak{p}}(x)$ is negative enough).
Hence, as above, the points with a very big "finite part" of the height of y_1 can be expressed, in each coordinate, by the Puiseux series described above.

After these observations we can finally proceed with the height bound which is the content of Lemma 5.4.2. We observe that it is sufficient to bound $h(x)$ in terms of $h(p_0), h(p_1), \ldots, h(p_s)$: each y_i, for $1 \leq i \leq s$, is an algebraic function of x and hence the height $h(y_i)$ is bounded above by $c \cdot (h(x) + 1)$ for a constant c (that can be chosen independent of i). We also point out that we restricted ourselves to the open affine $\mathbb{A}^s \subseteq \mathbb{P}_s$, but $C' \cap (\mathbb{P}_s \setminus \mathbb{A}^s)$ only consists of finitely many points, so our restriction is enough to conclude anyway.

- Let us start with the finite part. Let $(y_1, y_2, \ldots, y_s) \in (C' \cap \mathbb{A}^s) \cap L_{\underline{p}}$ be defined over $\overline{\mathbb{Q}}$ and let K'' be a number field containing x (for a Q'-th root of unity of y_1), all the y_i, for $1 \leq i \leq s$, and all the coefficients of all the Puiseux series expansions of the y_i, as described above.
  Let $\mathfrak{p}$ be a prime of K''; by Eisenstein's observation there are nonzero $\alpha_0, \alpha \in K''$ such that, for every Puiseux series expansion of every y_i

$$y_i = a_{r,i}x^r + a_{r-1,i}x^{r-1} + a_{r-2,i}x^{r-2} + \ldots$$

  then $\alpha_0 \cdot \alpha^j \cdot a_{r-j,i}$ is an integer of K'' for every natural number j. In particular, if $v_{\mathfrak{p}}(x) < -v_{\mathfrak{p}}(\alpha)$ then the Puiseux series converge.
  Let us then suppose $v_{\mathfrak{p}}(x) < -v_{\mathfrak{p}}(\alpha)$ and we express the condition $(y_1, y_2, \ldots, y_s) \in L_{\underline{p}}$ as

$$p_0 + p_1y_1 + \ldots + p_sy_s = 0$$

$$(p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s})x^r + (p_0a_{r-1,0} + p_1a_{r-1,1} + \ldots + p_sa_{r-1,s})x^{r-1} + \ldots = 0$$

  where we have set $a_{r-j,0} = 0$ unless $j = r$, in which case $a_{0,0} = 1$ (this corresponds to the constant Puiseux series 1). Let us suppose that $p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s}$ is nonzero; then the equality above implies that there is some positive integer $j > 0$ such that

$$(p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s})x^r$$

  and

$$(p_0a_{r-j,0} + p_1a_{r-j,1} + \ldots + p_sa_{r-j,s})x^{r-j}$$

  have the same valuation at $\mathfrak{p}$. If we let γ be $\min_{0 \leq k \leq s}\{v_{\mathfrak{p}}(p_i)\}$, then the valuation of the latter expression is bounded below by

$$(r-j)v_{\mathfrak{p}}(x) - jv_{\mathfrak{p}}(\alpha) - v_{\mathfrak{p}}(\alpha_0) + \gamma$$

by Eisenstein's observation and we therefore obtain

$$(r-j)v_{\mathfrak{p}}(x) - jv_{\mathfrak{p}}(\alpha) - v_{\mathfrak{p}}(\alpha_0) + \gamma \leq rv_{\mathfrak{p}}(x) + v_{\mathfrak{p}}(p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s})$$

which we rewrite as

$$-v_{\mathfrak{p}}(\alpha_0) + \gamma - v_{\mathfrak{p}}(p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s}) \leq j(v_{\mathfrak{p}}(x) + v_{\mathfrak{p}}(\alpha)) \leq v_{\mathfrak{p}}(x) + v_{\mathfrak{p}}(\alpha)$$

where the last inequality holds since $v_{\mathfrak{p}}(x) < -v_{\mathfrak{p}}(\alpha)$. We can rewrite the inequality as

$$-v_{\mathfrak{p}}(x) \leq v_{\mathfrak{p}}(\alpha) + v_{\mathfrak{p}}(\alpha_0) + \min_{0\leq k\leq s}\{v_{\mathfrak{p}}(p_k)\} + v_{\mathfrak{p}}(p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s}).$$

We can now finally estimate the finite part of the height x. Here comes a crucial observation: if we choose another prime lying above p, say $\mathfrak{p}'$, then the embedding $\overline{\mathbb{Q}} \subseteq \mathbb{C}_p$ has to be modified and, more worryingly, the branch of the Puisex series expansion of y_i we are considering might change. Thus, the same bound as above cannot hold just by replacing $\mathfrak{p}$ with $\mathfrak{p}'$.
Nevertheless, if the new branches have Puiseux series expansion

$$y_i = a'_{r,i}x^r + a'_{r-1,i}x^{r-1} + a'_{r-2,i}x^{r-2} + \ldots$$

for a suitable sequence $a'_{r-j,i}$ with $1 \leq i \leq s$ and $j > 0$ we obtain indeed

$$-v_{\mathfrak{p}'}(x) \leq v_{\mathfrak{p}'}(\alpha) + v_{\mathfrak{p}'}(\alpha_0) + \min_{0\leq k\leq s}\{v_{\mathfrak{p}'}(p_k)\} + v_{\mathfrak{p}'}(p_0a'_{r,0} + p_1a'_{r,1} + \ldots + p_sa'_{r,s})$$

while had assumed $p_0a'_{r,0} + p_1a'_{r,1} + \ldots + p_sa'_{r,s}$ nonzero. If all these values, for any choice of branches for each of the y_i, are nonzero, we can then bound

$$h_{fin}(x) \leq h(\alpha^{-1}) + h(\alpha_0^{-1}) + h(p_1) + h(p_2) + \ldots + h(p_s) + Z$$

where Z is the sum, over all the possible choices of branches for each y_i, of

$$h((p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s})^{-1})$$

which is a quantity equal to

$$h(p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s})$$

and therefore bounded by a suitable linear polynomial in $h(p_1), h(p_2), \ldots, h(p_s)$. This would conclude for the finite part, but we still need to deal with the case of $p_0a_{r,0} + p_1a_{r,1} + \ldots + p_sa_{r,s}$ equal to zero.

For any determination of the Puiseux series for y_i, we consider the infinite matrix

$$\begin{pmatrix} a_{r,0} & a_{r-1,0} & \ldots \\ a_{r,1} & a_{r-1,1} & \ldots \\ \ldots & \ldots & \ldots \\ a_{r,s} & a_{r-1,s} & \ldots \end{pmatrix}$$

and we let G be the Grassmannian parameterising hyperplanes in $\mathbb{P}_s$, which is isomorphic to $\mathbb{P}_s$. Given a natural number k, we can see that the subset S_k of G consisting of the points $(p_0, p_1, \ldots, p_s)$ such that

$$p_0 a_{r-j,0} + p_1 a_{r-j,1} + \ldots + p_s a_{r-j,s} = 0$$

for every $0 \leq j \leq k$ is indeed a (possibly empty) linear subspace of $\mathbb{P}_s$.
Suppose that some $\underline{p} = (p_0, p_1, \ldots, p_s)$ is contained in S_k for every natural number k. Then $p_0 y_0 + p_1 y_1 + \ldots + p_s y_s = 0$ for infinitely many points of C' (this can be seen for instance plugging a very big x into the Puiseux series), which implies that $C \subseteq L_{\underline{p}}$.
In either case, there is a natural number M, such that $S_k = S_{k+1}$ whenever $k \geq M$.

Remark 5.5.2. *We point out that this very last assertion is not effective.*

Since each y_i has finitely many branches, M can be chosen to work independently of the Puiseux series expansion considered. This implies that, unless $C \subseteq L_{\underline{p}}$, there is some integer $r - M \leq r' \leq r$ such that $p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}$ is nonzero; as above we obtain that there is some positive integer $j > 0$ such that

$$(p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}) x^{r'}$$

and

$$(p_0 a_{r'-j,0} + p_1 a_{r'-j,1} + \ldots + p_s a_{r'-j,s}) x^{r'-j}$$

have the same valuation at $\mathfrak{p}$. The argument runs exactly as before: we let γ be $\min_{0 \leq k \leq s}\{v_{\mathfrak{p}}(p_i)\}$ and we obtain

$$-v_{\mathfrak{p}}(x) \leq v_{\mathfrak{p}}(\alpha) + v_{\mathfrak{p}}(\alpha_0) + \min_{0 \leq k \leq s}\{v_{\mathfrak{p}}(p_i)\} + v_{\mathfrak{p}}(p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s})$$

with $p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}$ nonzero. Such a nonzero quantity exists for some value $r - M \leq r' \leq r$, so we obtain a final bound for the finite part of the height of x, which is

$$h_{fin}(x) \leq h(\alpha^{-1}) + h(\alpha_0^{-1}) + h(p_1) + h(p_2) + \ldots + h(p_s) + Z_0 + Z_1 + \ldots + Z_m$$

where Z_j is the sum, over all the possible choices of branches for each y_i, of

$$h((p_0 a_{r-j,0} + p_1 a_{r-j,1} + \ldots + p_s a_{r-j,s})^{-1})$$

whenever $p_0 a_{r-j,0} + p_1 a_{r-j,1} + \ldots + p_s a_{r-j,s}$ is nonzero; the height above is equal to

$$h(p_0 a_{r-j,0} + p_1 a_{r-j,1} + \ldots + p_s a_{r-j,s})$$

and finally we obtain the linear upper bound required.

- The argument for the infinite part is fairly similar. In particular, as we have just observed, there is a natural number M such that, for any choice of Puiseux series expansion, the quantity $p_0 a_{r-j,0} + p_1 a_{r-j,1} + \ldots + p_s a_{r-j,s}$ is nonzero for every $0 \leq j \leq M$.
  Let $(y_1, y_2, \ldots, y_s) \in (C' \cap \mathbb{A}^s) \cap L_{\underline{p}}$ be a point such that $|y_1|$ is big enough to make all the Puiseux series expansions (in an appropriate Q'-th root x of y_1) of the y_i for $1 \leq i \leq s$ converge.
  We wish to estimate the infinite part of the height of x. If some Galois conjugate x' of x does not make at least one Puiseux series converge, then $|x'|$ is smaller than some finite quantity depending on C' only (i.e. the radius of convergence of such Puiseux series). We therefore need to bound $|x'|$ for any Galois conjugate of x that makes all the Puiseux series converge and thus it is enough to give arguments for x only (in place of any other Galois conjugate).

  Let us fix Puiseux series expansions for each y_i

$$y_i = a_{r,i} x^r + a_{r-1,i} x^{r-1} + a_{r-2,i} x^{r-2} + \ldots$$

  and we suppose

$$p_0 + p_1 y_1 + \ldots + p_s y_s = 0$$

$$(p_0 a_{r,0} + p_1 a_{r,1} + \ldots + p_s a_{r,s}) x^r + (p_0 a_{r-1,0} + p_1 a_{r-1,1} + \ldots + p_s a_{r-1,s}) x^{r-1} + \ldots = 0$$

  where again $a_{r-j,0} = 0$ unless $j = r$, in which case $a_{0,0} = 1$. From the observations above, there is some $r - M \leq r' \leq r$ (where M is a constant depending on C' only) such that $p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}$ is nonzero. Moreover, choosing r' as the maximum such number, we can observe that, for some $j > 0$, the equality above implies

$$\left|(p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}) x^{r'}\right| \leq 2^j \left|(p_0 a_{r'-j,0} + p_1 a_{r'-j,1} + \ldots + p_s a_{r'-j,s}) x^{r'-j}\right|$$

  as, if this was not the case, then

$$\sum_{j=1}^{+\infty} \left|(p_0 a_{r'-j,0} + p_1 a_{r'-j,1} + \ldots + p_s a_{r'-j,s}) x^{r'-j}\right| < \left|(p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}) x^{r'}\right|.$$

  We now observe that there are constants β_0, β such that $|a_{r-j,i}| \leq \beta_0 \cdot \beta^j$ for every natural number j and every $0 \leq i \leq s$ (this is an immediate consequence of the fact that the Puiseux series obtained by algebraic functions converge, i.e. the Newton-Puiseux Theorem). Thus, we obtain

$$\left|(p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s}) x^{r'}\right| \leq 2^j \cdot (|p_0| + |p_1| + \ldots + |p_s|) \cdot \beta_0 \cdot \beta^{j+r-r'} \cdot |x|^{r'-j}$$

  and hence

$$|x|^j \leq (2\beta)^j \cdot (|p_0| + |p_1| + \ldots + |p_s|) \cdot \beta_0 \cdot \beta^M \cdot |(p_0 a_{r',0} + p_1 a_{r',1} + \ldots + p_s a_{r',s})|^{-1}.$$

Thus the infinite part of x can be bounded as

$$h_{inf}(x) \leq h(2\beta) + h(p_1) + h(p_2) + \ldots + h(p_s) + h(\beta_0 \cdot \beta^M) + Z_0 + Z_1 + \ldots + Z_M$$

where Z_j is the sum of $h(p_0 a_{r-j,0} + p_1 a_{r-j,1} + \ldots + p_s a_{r-j,s})$ for all the possible determinations of the branches of the y_i for $0 \leq i \leq s$. All the quantities involved are bounded linearly in $h(p_0), h(p_1), \ldots, h(p_s)$ and thus our Lemma is finally proven.

5.6 Final arguments

We now conclude the proof of Theorem 5.3.4.

Let K be a number field, let $C \subseteq \mathbb{G}_{\mathrm{m}}^n$ be a curve defined over K and let $\underline{c} = (c_1, \ldots, c_m)$ be a vector of multiplicatively independent nonzero elements of K. We consider any point $P \in C$ such that there is some $\underline{c}$-special hypersurface U such that C is tangent to U at P. Let us assume that U is contained in the zero locus of

$$c_1^{q_1} \ldots c_m^{q_m} \cdot x_1^{p_1} \ldots x_n^{p_n} = \omega$$

for integers $p_1, \ldots, p_n$ and integers $q_1, \ldots, q_m$ such that not all of the p_i are zero and such that

$$\gcd(p_1, \ldots, p_n, q_1, \ldots, q_m) = 1$$

and ω a root of unity.

Let Γ be the subgroup of $\mathbb{Z}^{m+n}$ consisting of the integer vectors $(a_1, \ldots, a_n, b_1, \ldots, b_m)$ such that

$$c_1^{b_1} \ldots c_m^{b_m} \cdot x_1(P)^{a_1} \ldots x_n(P)^{a_n} = 1.$$

We observe that the rank of Γ is at least one, since $P \in U$ (and not all of the p_i are zero). There are two possibilities:

1. if the rank of Γ is greater or equal than two, there are finitely many possible values as a consequence of Zilber-Pink for a curve in a power of $\mathbb{G}_{\mathrm{m}}$, which has been proven by a combination of the work of Bombieri-Masser-Zannier [10] and Maurin [37]. Namely, let us consider $\mathbb{G}_{\mathrm{m}}^m \times \mathbb{G}_{\mathrm{m}}^n$ and the map:

$$\iota : \mathbb{G}_{\mathrm{m}}^n \to \mathbb{G}_{\mathrm{m}}^m \times \mathbb{G}_{\mathrm{m}}^n;$$

$$\iota(x_1, \ldots, x_n) = (c_1, \ldots, c_m, x_1, \ldots, x_n).$$

This is just the same ι as above, with $n - m$ replaced by n. Observe that $\iota(C)$ is a curve.

Lemma 5.6.1. *$\iota(C)$ is not contained in any proper special subvariety.*

Proof. Suppose that $\iota(C) \subseteq H$ for some irreducible component of the zero locus of

$$y_1^{r_1} \dots y_m^{r_m} x_1^{t_1} \dots x_n^{t_n} = 1$$

with r_j, for $1 \leq j \leq m$, and t_i, for $1 \leq i \leq n$, integers, not all of them equal to zero; here $y_1, \dots, y_m$ and $x_1, \dots, x_n$ are the coordinates of $\mathbb{G}_{\mathrm{m}}^m$ and $\mathbb{G}_{\mathrm{m}}^n$ respectively. This implies that $C \subseteq \mathbb{G}_{\mathrm{m}}^n$ is contained in the zero locus of

$$c_1^{r_1} \dots c_m^{r_m} x_1^{t_1} \dots x_n^{t_n} = 1$$

which is a finite union of weakly special subvarieties. Since C is not contained in any proper such, this implies that all the t_i are zero, but then

$$c_1^{r_1} \dots c_m^{r_m} = 1$$

which implies that all the r_j are zero, as $c_1, \dots, c_m$ are multiplicatively independent. □

This Lemma allows us to apply the quoted result of Maurin [37], which is Zilber-Pink for a curve in a power of the multiplicative group. Such result implies that the values of $\iota(P)$ that correspond to a Γ of rank greater or equal than two are finitely many; this yields finiteness for the possibilities of P under the same rank constraint.

2. Γ has rank one; namely, it is generated by a vector $(v_1, \dots, v_n, w_1, \dots, w_m) \in \mathbb{Z}^{n+m}$. We use a result of Masser [38] that implies the following inequality, for an absolute constant $\eta_{n+m} > 0$:

$$\max\{|v_1|, \dots, |v_n|, |w_1|, \dots, |w_m|\} \leq$$
$$\leq \eta_{m+n}[K(P) : \mathbb{Q}]^{m+n+1} \max\{h(c_1), \dots, h(c_m), h(x(P_1)), \dots, h(x(P_n))\}^{m+n-1}.$$

We observe that, as $(p_1, \dots, p_n, q_1, \dots, q_m)$ is indivisible in $\mathbb{Z}^{n+m}$ (i.e. its entries are coprime), from

$$c_1^{q_1} \dots c_m^{q_m} \cdot x_1(P)^{p_1} \dots x_n(P)^{p_n}$$

being a root of unity, we obtain that $(v_1, \dots, v_n, w_1, \dots, w_m)$ is a multiple of $(p_1, \dots, p_n, q_1, \dots, q_m)$, hence

$$\max\{|p_1|, \dots, |p_n|, |q_1|, \dots, |q_m|\} \leq \max\{|v_1|, \dots, |v_n|, |w_1|, \dots, |w_m|\}.$$

We have an absolute upper bound on $[K(P) : \mathbb{Q}]$ by Proposition 5.2.1 and combining the two inequalities we obtain

$$\max\{|p_1|, \dots, |p_n|, |q_1|, \dots, |q_m|\} \leq \rho(1 + h(P))^{m+n-1}$$

where $\rho > 0$ is a constant depending on C (which determines $c_1, \dots, c_m$) only. But Proposition 5.4.1 gives a bound:

$$h(P) \leq \rho'(1 + \log \max\{|p_1|, \dots, |p_n|\})$$

for a constant $\rho' > 0$ depending on C only. Combining these we obtain an upper bound on $|p_1|, \dots, |p_n|$ and hence an upper bound on $h(P)$; together with the upper bound for $[K(P) : K]$, we conclude the proof of Theorem 5.3.4 by Northcott's principle.

Chapter 6

Tangencies in powers of modular curves

In this chapter we discuss a new kind of "double unlikely intersections" problem, taking place in $Y(1)^2$.

Before our discussion, we quote the result of Corvaja-Demeio-Masser-Zannier in [14]. We consider the elliptic scheme

$$\mathcal{E} : \{(x, y, \lambda) \in \mathbb{A}^3 \mid y^2 = x(x-1)(x-\lambda)\}$$

parameterised by $\lambda \in \mathbb{A}^1 \setminus \{0, 1\}$, which has the property that every fiber E_λ, defined as the zero locus

$$\{(x, y) \in \mathbb{A}^2 \mid y^2 = x(x-1)(x-\lambda)\}$$

has the structure of an elliptic curve; we implicitly consider the projective closure as usual.

Theorem 6.0.1 (Corvaja-Demeio-Masser-Zannier)**.** *Let $\rho : C \to \mathbb{A}^1\{0,1\}$ be a finite cover, defined over $\overline{\mathbb{Q}}$, and let $\sigma_1, \ldots, \sigma_n : C \to \mathcal{E}$ be algebraic sections of the elliptic scheme $\mathcal{E} \to \mathbb{A}^1\{0,1\}$ and let Γ be the group generated by those sections. There are only finitely many points $P \in C(\overline{\mathbb{Q}})$ for which there exist a positive integer n and any nontorsion section $\sigma \in \Gamma$ such that:*

1. *$\sigma(P)$ is an n-torsion point of $E_{\rho(P)}$;*
2. *$\sigma(C)$ is tangent to the torsion divisor $\mathcal{E}[n]$ at $\sigma(P)$.*

The Theorem requires rather different arguments than those involved in the proof of Theorem 5.0.1: indeed, the main step in the Corvaja-Demeio-Masser-Zannier result is essentially an o-minimality argument.

We now introduce the main result of this chapter, which is a "double unlikely intersections" result in $Y(1)^2$.
The situation is different from both the multiplicative or the (relative) abelian cases.
Indeed, the geometry of $Y(1)$ is not as simple as that of $\mathbb{G}_{\mathrm{m}}$, since the j-invariant satisfies a

"much worse" differential equation than exp; we refer to Remark 6.0.3 right below.
Moreover, the main obstacle in our proof is the necessity of an upper bound on the height: while in the multiplicative context the results by Bombieri-Masser-Zannier [10] and Habegger [27] and in the abelian context the result by Habegger [26] generally allow to bound heights from above, none of these results is available for $Y(1)$. We need to proceed directly with explicit calculations on the Puiseux series.
Before giving more details, we state the Theorem.

Theorem 6.0.2. *Let $C \subseteq Y(1)^2$ be a non-special curve over $\overline{\mathbb{Q}}$. There are only finitely many $P \in C(\overline{\mathbb{Q}})$ such that C is tangent to a one-dimensional special subvariety of $Y(1)^2$ at P.*

Remark 6.0.3. *We point out, as anticipated above, that, as in the case of Theorem 5.0.1, our Theorem 6.0.2 becomes much simpler if we ask for "quadruple intersection" rather than "double" ones.*

Remark 6.0.4. *The corresponding statement in higher dimensions is an immediate consequence of our explicit description of section 3.8: any one-codimensional special subvariety is given by the vanishing of a single modular polynomial.*
Therefore, for any curve C in $Y(1)^r$, there are only finitely many points on C at which C is tangent to a one-codimensional special subvariety of $Y(1)^r$, unless C is contained in a special subvariety of $Y(1)^r$.

As for Theorem 5.0.1, we can also in this case get an "avoiding problem" statement from the proof of Theorem 6.0.2.
Given a "continuous family" $\mathcal{F}$ of curves of $Y(1)^2$ (i.e. a subvariety of an appropriate Chow variety), we denote with T_n the set of curves $C \in \mathcal{F}$ such that C is tangent to the modular curve Y_n at some point.

Proposition 6.0.5. *Let us suppose that, for every positive integer n,*

$$T_n \text{ is not Zariski-dense in } \mathcal{F}.$$

Then, for "most" choices of $C \in \mathcal{F}$, there is no positive integer n such that C is tangent to Y_n at any point.

The hypothesis on the T_n is observed to be necessary: we can consider any family which is "identically tangent" to a single modular curve Y_n, e.g. the family of lines ℓ_P parameterised by $P \in Y_n$ and outlined by the requirement of ℓ_P being tangent to Y_n at P.

The proof is rather similar to that of Theorem 6.0.2. The main difference, exactly as in the previous chapter, lies in the fact that it is necessary for us to keep track of the growth of the coefficients of the Puiseux series attached to the branches of $C \in \mathcal{F}$, namely obtaining an upper bound on the constants c_0 and c as discussed in section 5.
We refer to section 5.5 for the basic notions on Puiseux series expansions and the meaning of

c and c_0.

Let us start diving into the proof of Theorem 6.0.2. We sketch the main points of the argument.
Let $P \in C(\overline{\mathbb{Q}})$ a point at which C is tangent to the modular curve Y_n.

- We start using the Pila-Wilkie Theorem 2.2.5, in order to bound from above the degree of P in terms on n. The main difference with our previous applications now lies on the presence of derivatives. We use a version of the Ax-Schanuel for the j-function Theorem 3.4.9 with derivatives by Pila-Tsimerman.

- We need to prove a form of weakly bounded height for $h(P)$. We can bound the infinite part of the height investigating the image $j^{-1}(C)$ in D^2 and obtaining some approximations that contradict Linear Forms in Logarithms, i.e. Baker's Theorem of [7].

- The most challenging part is the upper bound on the finite part of the height. This follows from a "bootstrap" argument on the coefficients of the Puiseux series expansions of the branches of C and Y_n. We first sketch the proof with some additional assumptions that make our argument easier. The "bootstrap" reasoning implies that C is the image via j of an algebraic variety and by functional transcendence we obtain that C is special. This very last part of the proof is not effective.

We start with a familiar argument and we use the Pila-Wilkie Theorem 2.2.5 together with the Ax-Schanuel for the j-function Theorem 3.4.9 with derivatives by Pila-Tsimerman.

We wish to construct a suitable definable set parameterising the entries of a matrix representing a "tangent isogeny". Let us first notice that the derivative of the upper half-plane transformation

$$z \to \frac{az+b}{cz+d}$$

is given by

$$z \to \frac{ad-bc}{(cz+d)^2}.$$

We denote with $f_1(t), f_2(t), \ldots \in \overline{\overline{\mathbb{Q}}(t)}$ a (definable) collection of finitely many branches of C, with the property that for all but finitely many $(x,y) \in C$ there is some $f_i(t)$ such that

$$y = f_i(x).$$

We can therefore construct the definable set (where j is intended restricted to D as usual)

$$S = \{(a,b,c,d) \in \mathbb{R}^4 \mid \exists x \in C(\mathbb{C}) \text{ in the domain of a branch } f$$

$$\text{such that } \frac{aj^{-1}(x)+b}{cj^{-1}(x)+d} = j^{-1}(f(x)) \text{ and } \frac{1}{(cj^{-1}(x)+d)^2} \cdot (j^{-1})'(x) = (j^{-1})'(f(x)) \cdot f'(x)\}.$$

In this case we do not only need to take into account integral points of S, since we are replacing $ad-bc$ by 1: indeed, if an isogeny corresponds to the integer matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

then, denoted by $M = \sqrt{AD-BC}$, the matrix with coefficients

$$\begin{pmatrix} A/M & B/M \\ C/M & D/M \end{pmatrix}$$

represents the same isogeny and it has moreover determinant equal to 1. Thus, any "tangent isogeny" of degree n gives rise to a point of degree 2 and exponential height bounded above by n, by the usual application of Lemma 2.5.12.

We now prove that S does not contain any connected semialgebraic curve. Let us suppose that there is a semialgebraic curve contained in S given by the coordinate functions $a(t), b(t), c(t), d(t)$, defined on the interval I that, as usual, we enlarge to the complex open neighborhood U; we refer to section 3.4 for the details.
We point out that $a(t), b(t), c(t), d(t)$ are extended to algebraic functions of t and that we can choose an holomorphic function $x(t)$ (by the nice properties of o-minimal structure, for instance "Definable Choice", that are explained in the notes of Starchenko [62]) satisfying identically

$$\frac{a(t)j^{-1}(x(t))+b(t)}{c(t)j^{-1}(x(t))+d(t)} = j^{-1}(f(x(t))) \text{ and}$$

$$\frac{1}{(c(t)j^{-1}(x(t))+d(t))^2} \cdot (j^{-1})'(x(t)) = (j^{-1})'(f(x(t))) \cdot f'(x(t)).$$

This property contradicts a form of Ax-Schanuel type Theorem for the j-function, this time involving also derivatives. This has been proven by Pila-Tsimerman in [54].

Theorem 6.0.6 (Pila-Tsimerman, Ax-Schanuel with derivatives)**.** *Let $u_1(t), \dots, u_n(t)$ be non-constant holomorphic functions defined on a connected open subset $U \subseteq \mathbb{C}$, such that $u_i(t) \in \mathbb{H}$ for every index i and every $t \in U$. Suppose that for no positive integer m and indices $1 \leq p, q \leq n$ then $\Phi_m(j(u_p(t)), j(u_q(t))) = 0$ identically. Then the field*

$$\mathbb{C}(t, u_1(t), \dots, u_n(t), j(u_1(t)), \dots, j(u_n(t)), j'(u_1(t)), \dots, j'(u_n(t)))$$

has transcendence degree at least $2n$ over $\mathbb{C}(t)$.

We apply the Theorem with $n = 2$; we set

$$u_1(t) = j^{-1}(x(t)) \text{ and } u_2(t) = j^{-1}(f(x(t)))$$

and, noting that

$$(j^{-1})'(s) = \frac{1}{j'(j^{-1}(s))} \text{ for any } s \in Y(1)$$

we can rewrite the equations above, defining S, as

$$\frac{a(t)u_1(t)+b(t)}{c(t)u_1(t)+d(t)} = u_2(t) \text{ and}$$

$$\frac{1}{(c(t)u_1(t)+d(t))^2}\cdot\frac{1}{j'(u_1(t))} = \frac{1}{j'(u_2(t))}\cdot f'(j(u_1(t)))$$

together with the implicit

$$j(u_2(t)) = f(j(u_1(t)) \text{ or } (j(u_1(t)), j(u_2(t))) \in C(\mathbb{C})$$

from which we deduce that the transcendence degree over $\mathbb{C}(t)$ of the field

$$\mathbb{C}(t, u_1(t), u_2(t), j(u_1(t)), j(u_2(t)), j'(u_1(t)), j'(u_2(t)))$$

is at most 3, since $u_1(t)$ and $u_2(t)$ (as well as $j(u_1(t))$ and $j(u_2(t))$) are algebraically dependent over $\mathbb{C}(t)$, while $j'(u_1(t))$ and $j'(u_2(t))$ are dependent over $\mathbb{C}(t, u_1(t), u_2(t), j(u_1(t)), j(u_2(t)))$. We directly obtain that C is special, from which we conclude.

We can therefore apply the Pila-Wilkie Theorem 2.2.5. Let P be a point of double intersection of C and the modular curve Y_n. Let d be the degree

$$[\mathbb{Q}(P) : \mathbb{Q}(C)]$$

and, by taking the d conjugates of P, we produce d points of degree 2 and exponential height bounded above by n on S. The Pila-Wilkie Theorem implies that

for every $\varepsilon > 0$, there is some $c > 0$, depending on ε only, such that $d < c \cdot n^{\varepsilon}$.

In order to conclude the proof, we also recall the isogeny estimates Theorem 3.2.2 and we obtain

$$n < c' \cdot (h(P) \cdot d)^{2.1}$$

for some explicit universal c', at least when P is not a special point; we already know by the André-Oort Conjecture (proven in [53]) that only finitely many special points lie on C.
Thus, we just need to prove a form of weakly bounded height:

$$h(P) < c'' \cdot \log n$$

for some positive constant c'' depending on C only.

6.1 Heights: infinite part

We now prove our weakly bounded height estimate for infinite places; namely, we prove suitable upper bounds on $|x|$ whenever $(x, y) \in C(\mathbb{C})$ is a point where C is tangent to Y_n.

We consider the Puiseux series expansion for a branch of C

$$y = b_0 x^{u/v} + b_1 x^{(u-1)/v} + b_2 x^{(u-2)/v} + \dots$$

where v is a positive integer, u is an integer and $b_0, b_1, b_2, \ldots$ is a succession of algebraic numbers satisfying

$$|b_i| < c_0 \cdot c^i$$

for some positive constants c_0 and c; we refer to section 5.5 in the previous chapter.
We wish to conjugate our Puiseux series via j in order to "see" it on D. Let $z \in D$ such that $\operatorname{Im} z$ is sufficiently big, i.e. such that $j(z)$ is in the radius of convergence at ∞ of the Puiseux series.
The q-expansion of the j-function

$$j(q) = \frac{1}{q} + 744 + 196884q + \ldots$$

maps holomorphically the unit disk onto $Y(1)$ and the quantity q is implicitly set equal to

$$q = \exp(2\pi i\tau)$$

and, in this sense, j (as a function of τ) maps bijectively D into $Y(1)$. We approximate $j^{-1}(y)$ using

$$j^{-1}(b_0x^{u/v} + \ldots) = \frac{1}{2\pi i}\log\left(b_0x^{-u/v} + O(x^{(-u-1)/v})\right) = \frac{\log b_0}{2\pi i} - \frac{u}{v}\frac{\log x}{2\pi i} + \log(1 + O(x^{-1/v}))$$

as long as $u > 0$; we write $x = j(z)$ and we use the approximations

$$x = j(z) = \exp(-2\pi iz) + 744 + O(|\exp(2\pi iz)|) = \exp(-2\pi iz) + O(1)$$

from which we deduce

$$\log x = -2\pi iz + \log(1 + O(|\exp(2\pi iz)|)) = -2\pi iz + O(\exp(-2\pi i \operatorname{Im} z)).$$

Moreover,

$$\log(1 + O(\left|x^{-1/v}\right|)) = O(\left|x^{-1/v}\right|) = e^{-s\operatorname{Im} z}$$

for some fixed $s > 0$ depending on C (and its Puiseux series expansions) only. Thus, we can finally approximate

$$j^{-1}(y) = px + r + O(e^{-s\operatorname{Im} z})$$

which, we recall, holds whenever $u > 0$.

We obtain a similar estimate when $u \leq 0$; in this case, the value $j^{-1}(z)$ will not "tend to infinity as x tends to infinity". Indeed, if $y \in Y(1)(\mathbb{C})$, there is a holomorphic function ξ, defined on an open neighborhood E of y, such that

$$j^{-1}(e) = \frac{1}{2\pi i}\log(\xi(e))$$

for every $e \in E$; we think of ξ as a section of the j-function from the unit disk to $Y(1)(\mathbb{C})$. We recall the expansion

$$y = b + O(\left|x^{-1/v}\right|)$$

and we finally obtain

$$j^{-1}(y) = \frac{1}{2\pi i}\log(\xi(b + O(|x^{-1/v}|))) = \frac{1}{2\pi i}\log(\xi(b)) + O(|x^{-1/v}|)$$

where we are crucially using that $\xi(b) \neq 0$. We point out that we are choosing only finitely many holomorphic functions ξ; namely, they only depend on b.
Comparing with our estimate for $|\log x|$ in terms of $\operatorname{Im} z$ obtained above, as long as $x = j(z)$ we obtain again

$$j^{-1}(y) = r + O(e^{-s\operatorname{Im} z})$$

for some s depending on C only.

We finally write down the condition of (x, y) belonging to the modular curve Y_n:

$$\frac{az+b}{cz+d} = pz + r + O(e^{-s\operatorname{Im} z}) \quad \text{(I)}$$

for some integers a, b, c, d satisfying $\gcd(a, b, c, d) = 1$ and $ad - bc = 1$. We observe that the derivatives must be equal by the "double intersection" requirement and we obtain

$$\frac{n}{(cz+d)^2} = p + O(e^{-s\operatorname{Im} z}) \qquad \text{(II)}$$

where we recall that z satisfies $j(z) = x$. We also point out that we safely ignore the big O signs while differentiating by the properties of holomorphic functions (namely, Cauchy's formulae) and that they imply constants which depend on C only.

Let us first suppose that $p \neq 0$. In this case, referring to the computations we performed above, $p = u/v$ (i.e. a positive rational number) and $2\pi i r$ is the logarithm of b_0, which is an algebraic number.
We suppose that

$$\operatorname{Im} z > \delta \cdot \log n$$

and we prove that, if δ is sufficiently big, we obtain a contradiction. From the fundamental Lemma 2.5.12 we have

$$|a|, |b|, |c|, |d| \leq n$$

and, from equation II, we deduce

$$c < O\left(\frac{n^{1/2}}{\delta \cdot \log n}\right)$$

as

$$|cz + d| \geq c \operatorname{Im} z.$$

We now expand equation II and we look at the real and imaginary parts:

$$n = p(c^2z^2 + 2cdz + d^2) + O(e^{-s\operatorname{Im} z})$$

$$n = p(c^2(x^2 - y^2) + 2cdx + d^2) + O(e^{-s\operatorname{Im} z})$$

$$0 = p(2xyc^2 + 2cdy) + O(e^{-s\,\mathrm{Im}\,z})$$

where we let $z = x + iy$, with x and y reals. From the very last equation, we obtain, unless $c = 0$,

$$-xc = d + O(e^{-s\,\mathrm{Im}\,z})$$

since $y \geq \frac{\sqrt{3}}{2}$. This in particular implies

$$|d| = \frac{1}{2}\,|c| + O(1)$$

as $x \in [-1/2, 1/2]$. This indeed contradicts

$$n = p(c^2(x^2 - y^2) + 2cdx + d^2) + O(e^{-s\,\mathrm{Im}\,z})$$

which, by $|x| \leq 1/2$ and the previous estimate on $|d|$, implies (by positivity of p)

$$n/p + c^2y^2 \leq c^2(1/4 + 1/2 + 1/4) + O(e^{-s\,\mathrm{Im}\,z}) \leq O\left(\frac{n}{\delta^2 \cdot (\log n)^2}\right) + O(e^{-s\,\mathrm{Im}\,z})$$

which is an immediate contradiction (with $c \neq 0$) as long as δ is sufficiently big.
Thus, $c = 0$ and from the above

$$ad = n = pd^2 + O(e^{-s\,\mathrm{Im}\,z})$$

$$\frac{a}{d} - p = O(e^{-s\,\mathrm{Im}\,z})$$

from which we deduce the equality $p = a/d$ by the following argument. Let $p = u/v$ for u, v positive integers; then

$$\left|\frac{a}{d} - \frac{u}{v}\right| \geq \frac{1}{|d|\,v}$$

unless $p = a/d$; since $|d|$ is bounded from above by n, we obtain, assuming $p \neq a/d$,

$$n^{-1} = O(e^{-s\,\mathrm{Im}\,z})$$

which is false as long as δ is big enough, i.e. $\mathrm{Im}\,z = \delta \cdot \log n$ is bigger than $s^{-1} \cdot \log n$.

We are almost done; we have proved $c = 0$ and $a/d = p$. We are left with the equation

$$b/d = r + O(e^{-s\,\mathrm{Im}\,z})$$

and we recall that $2\pi ir$ is the logarithm of an algebraic number. This prompts us to write, for an appropriate γ

$$\left|\log\gamma - \frac{b}{d}2\pi i\right| \leq O(e^{-s\,\mathrm{Im}\,z})$$

which, unless γ is a root of unity, by taking δ sufficiently big, contradicts Linear Forms in Logarithms, a transcendence result by Baker; we refer to Theorem 3.1 of [7].
In the latter case, then r is a rational number and, by the exact same argument as the one employed with p, we obtain $b/d = r$.

Thus, the value of the tuple (a, b, c, d) is uniquely determined up to scalar multiplication and hence, as long as δ is sufficiently big, there can be at most one modular curve Y_n which is tangent to a point of $C(\mathbb{C})$ satisfying $\operatorname{Im} z > \delta \cdot \log n$. This exception is essentially inevitable, since C could be a priori Y_n itself and therefore C would be tangent to Y_n at all of its points. Summarising, we have proved the following:

there are a positive integer N and a value of δ such that, if $(x, y) \in C(\mathbb{C})$ satisfies the "double intersection" condition with the modular curve Y_n, then either (x, y) is a point of the "unique exceptional modular curve" $Y_N(\mathbb{C})$ or

$$\log |x| \leq \delta \cdot \log n$$

where we are implicitly using the estimate above $\log |x| = O(\operatorname{Im} z)$.
This concludes the proof for the infinite part of the height whenever $p \neq 0$, as the average of $\log |\cdot|$ among the Galois conjugates of x will satisfy a like bound; we notice that, even if C might not be defined over $\mathbb{Q}$, the Galois conjugates of the curve will still be tangent to the very same modular curve (i.e. each Y_n is indeed defined over $\mathbb{Q}$).

The case $p = 0$ is easier; from

$$\frac{n}{(cz + d)^2} = O(e^{-s \operatorname{Im} z})$$

we deduce, using $|c|, |d| \leq n$

$$n \leq n^2 \cdot (\operatorname{Im} z)^2 \cdot O(e^{-s \operatorname{Im} z})$$

which implies

$$s \operatorname{Im} z - 2 \log (\operatorname{Im} z) \leq \log n$$

and in turn we obtain the correct bound

$$\operatorname{Im} z \leq \delta \cdot \log n.$$

This completes the proof for the infinite part of the height.

6.2 Heights: finite part (sketch)

In this section we sketch the proof of the weakly bounded height bound, for what concerns the finite part.
Our simplified proof mainly consists in prescribing a nice shape of the Puiseux series expansions, rather than working in full generality, as in the next section, and considering only "big" primes.

We consider a prime ideal $\mathfrak{p}$, lying above a "big" prime number p of $\mathbb{Q}$. Let us write a Puiseux series expansion for a branch of C

$$y = \alpha x^{u/v} + \beta x^{(u-1)/v} + \dots.$$

We assume, for now, that $u = v = 1$.
The Puiseux series expansion of the branch of the modular curve corresponding to the transformation of the upper half-plane $\mathbb{H}$

$$\tau \to \frac{a}{d}\,\tau + \frac{c}{d}$$

turns out to be of the shape

$$y = \omega x^{a/d} + \dots$$

where ω is a root of unity of order $d/(\gcd(c,d))$.
Since we assumed p sufficiently big, the series above converge whenever

$$v_{\mathfrak{p}}(x) < 0$$

and comparing the two expansions (with some extra assumptions on the Puiseux series) implies

$$v_{\mathfrak{p}}(x^{a/d}) = v_{\mathfrak{p}}(x)$$

from which we can obtain $a = d$.
Let us rename $d = n$ for convenience and hence, given the transformation

$$\tau \to \tau + \frac{c}{n}$$

the corresponding Puiseux series is

$$y = \omega x + 744(\omega - 1) + 196884(\omega^{-1} - \omega)x^{-1} + \dots$$

for ω a root of unity of order n (assuming c and n relatively prime).

We are therefore comparing Puiseux series of the following form:

$$y = \omega x + a_1(\omega) + a_2(\omega)x^{-1} + a_3(\omega)x^{-2} + \dots$$

$$y = \alpha x + b_1 + b_2 x^{-1} + b_3 x^{-2} + \dots$$

for ω a root of unity of order n, $a_i(t)$ fixed elements of $\mathbb{Z}[t, t^{-1}]$, α and b_i fixed algebraic numbers in the same number field: since p is big we assume that $\mathfrak{p}$ never occurs in their prime ideal factorisation with a negative exponent - this follows from the properties of Puiseux series of algebraic functions that we have seen already seen in section 5.5.
We use the tangency hypothesis and, taking a combination of the derivatives, we obtain

$$y - xy' = a_1(\omega) + 2a_2(\omega)x^{-1} + a_3(\omega)x^{-2} + \dots;$$

$$y - xy' = b_1 + 2b_2 x^{-1} + 3b_3 x^{-2} + \dots.$$

Comparing the equations for y we obtain

$$v_{\mathfrak{p}}(\omega - \alpha) \geq -v_{\mathfrak{p}}(x)$$

and, for $y - xy'$,

$$v_{\mathfrak{p}}(a_1(\omega) - b_1) \geq -v_{\mathfrak{p}}(x).$$

We have that ω and α coincide "modulo $\mathfrak{p}$" and thus we can substitute α instead of ω into the Laurent polynomial $a_1(t) \in \mathbb{Z}[t, t^{-1}]$.
Let us suppose that $b_1 \neq a_1(\alpha)$ and we obtain, forgetting about small primes,

$$h_{fin}(x) \leq h_{fin}((a_1(\alpha) - b_1)^{-1}).$$

The quantity on the right is a constant independent of n and this proves bounded height.

Remark 6.2.1. *We stress that the analogous estimate*

$$h_{fin}(x) \leq h_{fin}((\omega - \alpha)^{-1})$$

could not be deduced from the above. Indeed, we are implicitly considering all the Galois conjugates of $\mathfrak{p}$ while bounding $h_{fin}(x)$, but taking a Galois conjugate of $\mathfrak{p}$ might need us to consider a Galois conjugate of ω as well, since the embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}_p$ would change; moreover, different Galois conjugates might lie on different branches.

Therefore we can assume $b_1 = a_1(\omega)$. We replace b_1 in the equations above

$$y = \omega x + a_1(\omega) + a_2(\omega)x^{-1} + a_3(\omega)x^{-2} + \dots$$

$$y = \alpha x + a_1(\alpha) + b_2 x^{-1} + b_3 x^{-2} + \dots$$

and

$$y - xy' = a_1(\omega) + 2a_2(\omega)x^{-1} + a_3(\omega)x^{-2} + \dots;$$

$$y - xy' = a_1(\alpha) + 2b_2 x^{-1} + 3b_3 x^{-2} + \dots.$$

Now we recall from above that

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha)) \geq -v_{\mathfrak{p}}(x)$$

and, from the equations for y,

$$v_{\mathfrak{p}}(\omega - \alpha) \geq -2v_{\mathfrak{p}}(x).$$

We observe that, as $a_1(t)$ is an element of $\mathbb{Z}[t, t^{-1}]$, we obtain

$$(a_1(\omega) - a_1(\alpha))/(\omega - \alpha) \in \mathbb{Z}[\omega, \omega^{-1}, \alpha, \alpha^{-1}]$$

and this implies, since p is big enough not to appear in the prime factorisation of α (with a negative exponent),

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha)) \geq -2v_{\mathfrak{p}}(x)$$

and from the equations for $y - xy'$ (as long as $p > 2$):

$$v_{\mathfrak{p}}(a_2(\omega) - b_2) \geq -v_{\mathfrak{p}}(x).$$

Again, we observe that if $a_2(\omega) - b_2 \neq 0$ this would imply

$$h_{fin}(x) \leq h_{fin}((a_2(\alpha) - b_2)^{-1})$$

which is the required height estimate. Hence, we can assume $b_2 = a_2(\alpha)$.
This process can be performed indefinitely. We show the next step.

We replace b_2 in the equations above

$$y = \omega x + a_1(\omega) + a_2(\omega)x^{-1} + a_3(\omega)x^{-2} + \dots$$

$$y = \alpha x + a_1(\alpha) + a_2(\alpha)x^{-1} + b_3x^{-2} + \dots$$

and

$$y - xy' = a_1(\omega) + 2a_2(\omega)x^{-1} + a_3(\omega)x^{-2} + \dots;$$

$$y - xy' = a_1(\alpha) + 2a_2(\alpha)x^{-1} + 3b_3x^{-2} + \dots.$$

From the equations for y

$$v_{\mathfrak{p}}(\omega - \alpha) \geq -3v_{\mathfrak{p}}(x)$$

and, as a_2 is a polynomial in t and t^{-1}, we obtain

$$(a_2(\omega) - a_2(\alpha))/(\omega - \alpha) \in \mathbb{Z}[\omega, \omega^{-1}, \alpha, \alpha^{-1}].$$

This implies, for p big enough,

$$v_{\mathfrak{p}}(a_2(\omega) - a_2(\alpha)) \geq -3v_{\mathfrak{p}}(x).$$

We also recall

$$(a_1(\omega) - a_1(\alpha))/(\omega - \alpha) \in \mathbb{Z}[\omega, \omega^{-1}, \alpha, \alpha^{-1}]$$

hence

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha)) \geq -3v_{\mathfrak{p}}(x)$$

and, from the equations for $y - xy'$ (as long as $p > 3$),

$$v_{\mathfrak{p}}(a_3(\omega) - b_3) \geq -v_{\mathfrak{p}}(x).$$

Again, we observe that if $a_3(\omega) - b_3 \neq 0$ this would imply

$$h_F(x) \leq h_F((a_3(\alpha) - b_3)^{-1})$$

and so $b_3 = a_3(\alpha)$ and so on...

We can therefore assume that

$$b_i = a_i(\alpha_i)$$

for every positive integer i. It can be proved, as a consequence of the Ax-Schanuel for the j-function Theorem 3.4.9 by Pila-Tsimerman, that C is then necessarily special. This concludes

the argument.

We point out that this proof is not effective: we do not have a criterion to find any i such that $b_i \neq a_i(\alpha_i)$.

The arguments above even bound $h_{fin}(x)$ absolutely, in the sense that $h_{fin}(x)$ is bounded above by a constant depending on C only (and not on n).
In the more general proof, where we consider small primes and more complicated branches of the modular curve, our upper bound is indeed of the form $c \cdot \log n$, for a constant c depending on C only.

6.3 Heights: finite part

We now prove that whenever $(x, y) \in C(\overline{\mathbb{Q}})$ satisfies the "double intersection" condition with the modular curve Y_n, then the finite part $h_{fin}(x)$ of the height of x is bounded above by $c \cdot \log n$ for an appropriate c depending on C only.

We bound each single place at a time and this is performed by embedding "suitably" any number field into $\mathbb{C}_p$, in the sense that if our place of interest corresponds to the prime ideal $\mathfrak{p}$ then $\mathfrak{p}$ is exactly the preimage of the elements of $\mathbb{C}_p$ of positive p-adic norm.

Let us briefly sketch the structure of the proof:

1. first, we describe the Puiseux series expansions for C and the modular curve Y_n, computing suitable bounds on the coefficients that ensure convergence;

2. if the branch of the modular curve is corresponding to the matrix

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

   we prove the correct weakly bounded height estimate except for one single value of a/d (that is in principle inevitable, since C might be the very Y_n);

3. we perform a "bootstrap" argument and we prove constraints, on the assumption that our weakly bounded height bound does not hold, on all the coefficients of the Puiseux series attached to the chosen branch of C;

4. we finally prove, using an ineffective functional transcendence statement (the usual Ax-Schanuel for the j-function by Pila-Tsimerman, i.e. Theorem 3.4.9) that the constraints obtained above force C to be special.

We can start the argument.

1. Let us fix a finite place $\mathfrak{p}$ above the prime p of $\mathbb{Q}$. We take a Puiseux series expansion of a branch of C:

$$y = \alpha_0 x^{u/v} + \alpha_1 x^{(u-1)/v} + \dots$$

where the coefficients, as seen in section 5.5, are algebraic numbers which are all contained in a fixed number field and satisfying

$$c_0 \cdot c^i \alpha_i \text{ are algebraic integers}$$

for appropriate positive integers c_0 and c. Thus, whenever

$$v_{\mathfrak{p}}(x) < -v v_{\mathfrak{p}}(c)$$

then the Puiseux series above converges at x in $\mathbb{C}_p$.

We now consider the Puiseux series expansion of the branches of the modular curve Y_n; if its representing matrix on the upper half-plane is

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

with a, b, d integers such that $ad = n$ and $\gcd(a, b, d) = 1$, where we are implicitly setting $c = 0$, since we are interested at the behaviour of "big" z. There is no need to prove this assertion and we show below that indeed the Puiseux series corresponding to these matrices suffice for the purpose.

Let us recall the q-expansion of the j-function:

$$j(\tau) = \frac{1}{q} + 744 + 196884q + \dots \text{ where } q = \exp(2\pi i \tau)$$

and therefore

$$j\left(\frac{a}{d}\tau + \frac{b}{d}\right) = \frac{\omega}{q^{a/d}} + 744 + \frac{196884}{\omega} q^{a/d} + \dots$$

for ω a root of unity of order $\gcd(b, d)$; more precisely, $\omega = \exp(-2\pi i b/d)$.
The Puiseux series expansion of the corresponding branch f satisfying

$$y = f(x) \text{ i.e. } j\left(\frac{a}{d}\tau + \frac{b}{d}\right) = f(j(\tau))$$

can therefore be seen to be equal to

$$y = \omega x^{a/d} + a_1(\omega) x^{(a-1)/d} + a_2(\omega) x^{(a-2)/d} + \dots$$

where the Laurent polynomials $a_1(t), a_2(t)$ are "fixed" elements of $\mathbb{Q}[t, t^{-1}]$, in the sense that they depend on a/d only.
The coefficients of the Puiseux series above are obtained solving the linear system

$$a_0 j(\tau)^{a/d} + a_1 j(\tau)^{(a-1)/d} + a_2 j(\tau)^{(a-2)/d} + \dots = j\left(\frac{a}{d}\tau + \frac{b}{d}\right)$$

which is "triangular", in the sense that, for each $j(\tau)^{(a-i)/d}$, there is a unique element where q is occurring with the least exponent, namely $q^{-(a-i)/d}$.
We can even "solve" the system for each coefficient of $j\left(\frac{a}{d}\tau + \frac{b}{d}\right)$ individually:

$$a'_0 j(\tau)^{-k\cdot a/d} + a'_1 j(\tau)^{-k\cdot a/d-1} + a'_2 j(\tau)^{-k\cdot a/d-2} + \ldots = \frac{A}{\omega^k} q^{k\cdot a/d}$$

where A is an integer k is an integer ≥ -1. We just need to argue for the a'_i. We observe that

- the coefficients of $j\left(\frac{a}{d}\tau + \frac{b}{d}\right)$ are elements of $\mathbb{Z}[\omega, \omega^{-1}]$, as it is apparent from the expression above;
- the coefficients of any power

  $$j(\tau)^{-k\cdot a/d-i} = c_0 q^{-k\cdot a/d-i} + c_1 q^{-k\cdot a/d-i-1} + \ldots$$

  satisfy $d^{2j} \cdot c_j \in \mathbb{Z}$ for any natural number j, as it can be directly seen from the binomial expansion of $(1+t)^{d'/d}$ for any integer d'.

Thus, we claim by induction that $a'_i \cdot d^{2i}$ is an element of $\mathbb{Z}[\omega, \omega^{-1}]$.
This is automatically true for a'_0; let us suppose it true for $a'_0, a'_1, \ldots, a_{i-1}$. The coefficients of any occurrence (while solving the system) of $q^{-(a-i)/d}$ have the property that, if multiplied by d^{2i}, they become elements of $\mathbb{Z}[\omega, \omega^{-1}]$, with the only possible exception of the coefficient a'_i. This is indeed not an exception and the claim follows.
This observation has a nice consequence: our Puiseux series converges (in x) whenever

$$v_{\mathfrak{p}}(x) < -2v_{\mathfrak{p}}(d)$$

and, explicitly, $d^{2i}a_i$ is an algebraic integer for every natural number i.
We make right now one additional observation that will turn out to be useful later; indeed, it can be observed by the expansion of $j(\tau)$ and $j\left(\frac{a}{d}\tau + \frac{b}{d}\right)$ that, as

$$j(\tau)^{a/d} = \frac{1}{q^{a/d}} + 744\frac{a}{d} \cdot \frac{1}{q^{(a-d)/d}} + \ldots$$

- if $a/d \geq 1$, then all the coefficients $a_1, a_2, \ldots, a_{d-1}$ are equal to zero, as the first nonvanishing coefficient beside the leading term corresponds to the monomial $744\frac{a}{d} \cdot \frac{1}{q^{(a-d)/d}}$;
- if $a/d < 1$, then we can only assert the vanishing of the coefficients $a_1, a_2, \ldots, a_{a-1}$, since $j\left(\frac{a}{d}\tau + \frac{b}{d}\right)$ contains a term of "degree" zero (namely, 744).

We can finally compare our two Puiseux series expansions:

$$y = \alpha_0 x^{u/v} + \alpha_1 x^{(u-1)/v} + \ldots$$

$$y = \omega x^{a/d} + a_1(\omega) x^{(a-1)/d} + a_2(\omega) x^{(a-2)/d} + \ldots$$

where the former converges whenever $v_\mathfrak{p}(x) < -vv_\mathfrak{p}(c)$ for a positive integer c depending on C only and the latter converges whenever $v_\mathfrak{p}(x) < -2v_\mathfrak{p}(d)$. We are tacitly fixing a number field containing the coefficients of the Puiseux series; this cannot be done once and for all since such number field needs to contain arbitrary roots of unity.
We stress the following important observation: for any point $(x, y) \in Y_n(\overline{\mathbb{Q}})$ with negative $v_\mathfrak{p}(x)$ we can choose an embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}_p$ and an appropriate branch whose Puiseux series expansion, when evaluated at x, converges to y; this is a direct consequence of the fact that the construction above produces exactly $\psi(n)$ matrices with distinct corresponding Puiseux expansions, which are all convergent whenever $v_\mathfrak{p}(x)$ is negative, and therefore all the admissible values of y corresponding to x are "reached" by the Puiseux series.

2. Our first goal is proving that $a/d = u/v$. Let us assume

$$v_\mathfrak{p}(x) < -vv_\mathfrak{p}(c) - 2v_\mathfrak{p}(d)$$

so that both the Puiseux series converge at x. We write

$$y = \alpha_0 x^{u/v} + \alpha_1 x^{(u-1)/v} + \dots$$

$$y = \omega x^{a/d} + a_1(\omega) x^{(a-1)/d} + a_2(\omega) x^{(a-2)/d} + \dots$$

from which we immediately observe, by the condition $d^{2i} a_i(\omega)$ integral, that

$$\frac{a}{d} v_\mathfrak{p}(x) = v_\mathfrak{p}(y) = \frac{u-j}{v} v_\mathfrak{p}(x) + v_\mathfrak{p}(\alpha_j)$$

for some natural number j. We can also differentiate both expressions

$$xy' = \frac{u}{v}\alpha_0 x^{u/v} + \frac{u-1}{v}\alpha_1 x^{(u-1)/v} + \dots$$

$$xy' = \frac{a}{d}\omega x^{a/d} + \frac{a-1}{d} a_1(\omega) x^{(a-1)/d} + \frac{a-2}{d} a_2(\omega) x^{(a-2)/d} + \dots$$

that we multiplied by x for convenience; these two expressions need to be equal at x by the "double isogeny" condition.
From the equations for y we obtain

$$v_\mathfrak{p}\left(\alpha_0 x^{u/v} - \omega x^{a/d}\right) \geq \min\{v_\mathfrak{p}(\alpha_i x^{(u-i)/v}), v_\mathfrak{p}(a_k(\omega) x^{(a-k)/d})\}$$

for some positive integers i and k (with k additionally satisfying $k \geq \min\{|a|, |d|\}$ by our observation above) and, likewise, from the equations for xy' we read

$$v_\mathfrak{p}\left(\frac{u}{v}\alpha_0 x^{u/v} - \frac{a}{d}\omega x^{a/d}\right) \geq \min\{v_\mathfrak{p}\left(\frac{u-i}{v}\alpha_i x^{(u-i)/v}\right), v_\mathfrak{p}\left(\frac{a-k}{d} a_k(\omega) x^{(a-k)/d}\right)\}$$

for suitable positive integers i and k (with k satisfying the same constraints). Thus, combining the two estimates, we can "cancel" ω:

$$v_\mathfrak{p}\left(\left(\frac{u}{v} - \frac{a}{d}\right)\alpha_0 x^{u/v}\right) \geq \min\{v_\mathfrak{p}\left(\frac{1}{vd}\alpha_i x^{(u-i)/v}\right), v_\mathfrak{p}\left(\frac{1}{d} a_k(\omega) x^{(a-k)/d}\right)\}$$

where we are using a "weaker" lower bound.
By assumption on x, we have both

$$v_{\mathfrak{p}}(\alpha_i \cdot x^{-i/v}) \geq -i\left(v_{\mathfrak{p}}(c) + \frac{1}{v}v_{\mathfrak{p}}(x)\right) - v_{\mathfrak{p}}(c_0)$$

and

$$v_{\mathfrak{p}}(a_k(\omega) \cdot x^{-k/d}) \geq -k\left(\frac{2}{d}v_{\mathfrak{p}}(d) + \frac{1}{d}v_{\mathfrak{p}}(x)\right)$$

from which finally we deduce

$$v_{\mathfrak{p}}\left(\left(\frac{u}{v} - \frac{a}{d}\right)\alpha_0 x^{u/v}\right) \geq -v_{\mathfrak{p}}(c_0) - v_{\mathfrak{p}}(v) - v_{\mathfrak{p}}(d)$$

$$+\min\{\frac{u}{v}v_{\mathfrak{p}}(x) - i\left(v_{\mathfrak{p}}(c) + \frac{1}{v}v_{\mathfrak{p}}(x)\right), \frac{a}{d}v_{\mathfrak{p}}(x) - k\left(\frac{2}{d}v_{\mathfrak{p}}(d) + \frac{1}{d}v_{\mathfrak{p}}(x)\right)\}$$

and we obtain either

$$\frac{u}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right) - v_{\mathfrak{p}}(\alpha_0) - v_{\mathfrak{p}}(c_0) - v_{\mathfrak{p}}(v) - v_{\mathfrak{p}}(d) + \frac{u}{v}v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(c) - \frac{1}{v}v_{\mathfrak{p}}(x)$$

from which we obtain

$$v_{\mathfrak{p}}(x) \geq -vv_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right) - v_{\mathfrak{p}}(d) - v_{\mathfrak{p}}(C_0)$$

where C_0 is a fixed positive integer depending on C only, or (when $a \geq d$)

$$\frac{u}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right) - v_{\mathfrak{p}}(\alpha_0) - v_{\mathfrak{p}}(c_0) - v_{\mathfrak{p}}(v) - v_{\mathfrak{p}}(d) + \frac{a}{d}v_{\mathfrak{p}}(x) - 2v_{\mathfrak{p}}(d) - v_{\mathfrak{p}}(x)$$

and, replacing the estimate obtained before, depending on a natural number j,

$$\frac{a}{d}v_{\mathfrak{p}}(x) = \frac{u-j}{v}v_{\mathfrak{p}}(x) + v_{\mathfrak{p}}(\alpha_j) \geq \frac{u}{v}v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(c_0)$$

we get

$$v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right) - 3v_{\mathfrak{p}}(d) - v_{\mathfrak{p}}(C_1)$$

where C_1 is a fixed positive integer depending on C only, or finally (when $a < d$)

$$\frac{u}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right) - v_{\mathfrak{p}}(\alpha_0) - v_{\mathfrak{p}}(c_0) - v_{\mathfrak{p}}(v) - v_{\mathfrak{p}}(d) + \frac{a}{d}v_{\mathfrak{p}}(x) - \frac{2a}{d}v_{\mathfrak{p}}(d) - \frac{a}{d}v_{\mathfrak{p}}(x)$$

and hence

$$\frac{u}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right) - 2v_{\mathfrak{p}}(d) - v_{\mathfrak{p}}(C_2)$$

where C_2 is a fixed positive integer depending on C only.

We are now left with the task of bounding from above $v_{\mathfrak{p}}\left(\frac{u}{v} - \frac{a}{d}\right)$. We first recall the expansions (truncated at the first term)

$$y = \alpha_0 x^{u/v} + \dots$$

$$y = \omega x^{a/d} + \dots$$

and we point out that there is a constant e_0 depending on C only such that whenever $v_{\mathfrak{p}}(x) < -e_o$ then

$$v_{\mathfrak{p}}(\alpha_0 x^{u/v}) = v_{\mathfrak{p}}(y) = v_{\mathfrak{p}}(\omega x^{a/d})$$

where the second equality occurs whenever the Puiseux series attached to Y_n converges, i.e. for

$$v_{\mathfrak{p}}(x) < -2v_{\mathfrak{p}}(d).$$

We obtain that, since we are assuming $a/d \neq u/v$,

$$v_{\mathfrak{p}}(\alpha_0) \neq 0$$

and therefore there are only finitely many choices (in the sense that these choices depend only on C) for the prime p lying below $\mathfrak{p}$.
In order to bound from above $v_{\mathfrak{p}}(ud - av)$, we observe that

$$|ud - av| \leq (|u| + |v|)n$$

and therefore, if we consider the product E_n of all the powers of the admissible primes (those who could be lying below $\mathfrak{p}$) bounded above by the quantity $(|u| + |v|)n$, by the reasoning outlined before, we obtain the bound

$$E_n \leq (|u| + |v|)^{e_1} n^{e_1}$$

for a constant e_1 depending on C only.
Summing up all of the above, in either case, unless $u/v = a/d$,

$$v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(E_n) - 3v_{\mathfrak{p}}(n) - v_{\mathfrak{p}}(e_1)$$

where e_2 is a positive integer depending on C only. In particular, ranging over all the admissible primes $\mathfrak{p}$ for the finite part of the height:

$$h_{fin}(x) \leq \log E_n + 3\log n + \log e_2 \leq \delta' + \delta \log n$$

for suitable constants δ, δ' depending on C only.

Remark 6.3.1. *We point out that, in this part, it was essential for us "getting rid" of ω: the embedding of $\overline{\mathbb{Q}}$ into $\mathbb{C}_p$ changes as long as we change $\mathfrak{p}$ and thus ω would need to be replaced by some Galois conjugate.*

Therefore, we just proved that, if $u/v \neq a/d$, the correct bound on the height holds for x.

3. We now show our "bootstrap" argument, that proves a suitable upper bound on the height. We recall that we are in the situation where $u/v = a/d$.
As we have observed before, the polynomials $a_1(t), a_2(t), \ldots$ occurring in the Puiseux

series expansions of the modular curves, depend on the ratio a/d only. Thus, we write again

$$y = \alpha_0 x^{u/v} + \alpha_1 x^{(u-1)/v} + \dots$$

$$y = \omega x^{u/v} + a_1(\omega) x^{(u-1)/v} + a_2(\omega) x^{(u-2)/v} + \dots$$

and a combination of its derivatives (where u is potentially allowed to be 0)

$$uy - vxy' = \alpha_1 x^{(u-1)/v} + 2\alpha_2 x^{(u-2)/v} + 3\alpha_3 x^{(u-3)/v} + \dots;$$

$$uy - vxy' = a_1(\omega) x^{(u-1)/v} + 2a_2(\omega) x^{(u-2)/v} + 3a_3(\omega) x^{(u-3)/v} + \dots.$$

We also observe that, once a/d is fixed (as in our situation), the value of n is roughly encoded in the order of the root of unity ω.

We assume again, in order to make all the Puiseux series converge at x, the condition

$$v_{\mathfrak{p}}(x) < -v v_{\mathfrak{p}}(c) - 2v_{\mathfrak{p}}(d) = -v_{\mathfrak{p}}(f_0)$$

and this time we could bound above the valuation of x by a fixed positive integer f_0 depending on C only. Moreover, we can also assume (enlarging f_0 in a way depending on C only), that

$$v_{\mathfrak{p}}(\alpha_0 x^{u/v}) < v_{\mathfrak{p}}(\alpha_i x^{(u-i)/v}) \text{ for every positive integer } i.$$

Comparing the equations for y we obtain

$$v_{\mathfrak{p}}((\omega - \alpha_0) x^{u/v}) \geq v_{\mathfrak{p}}(x^{(u-1)/v}) - v_{\mathfrak{p}}(G_0)$$

for a constant G_0 depending on C only, from which we obtain

$$\frac{1}{v} v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(\omega - \alpha_0) - v_{\mathfrak{p}}(G_0).$$

Remark 6.3.2. *We exploit the occasion to make sense of the following: we need to "get rid" of ω; indeed, we could not deduce the estimate*

$$h_{fin} \leq v h(\omega - \alpha_0) + h(G_0)$$

where the right-hand side is absolutely bounded (i.e. bounded above only in terms of C). If we wish to take into account the Galois conjugates of $\mathfrak{p}$, then we might need to conjugate ω as well (and, more in general, the Galois conjugates could lie on other branches of the relevant functions).

Comparing the equations for $uy - vxy'$ we obtain instead

$$v_{\mathfrak{p}}((a_1(\omega) - \alpha_1) x^{(u-1)/v}) \geq v_{\mathfrak{p}}(x^{(u-2)/v}) - v_{\mathfrak{p}}(H_0)$$

for a constant H_0 depending on C only, from which we obtain

$$\frac{1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_1(\omega) - \alpha_1) - v_{\mathfrak{p}}(H_0).$$

The key step at this point is replacing α_0 instead of ω into the polynomial $a_1(t) \in \overline{\mathbb{Q}}[t, t^{-1}]$. This can be done without altering the $\mathfrak{p}$-adic evaluation. We need to rewrite the estimate above as

$$v_{\mathfrak{p}}(\omega - \alpha_0) \geq -\frac{1}{v}v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(G_0) > 0$$

where the positivity assumption is ensured whenever $v_{\mathfrak{p}}(x)$ does not satisfy

$$-v_{\mathfrak{p}}(x) \leq v_{\mathfrak{p}}(I_0)$$

for a fixed (depending on C only) positive integer I_0. Since the values ω and α_0 differ by an element of $\mathbb{C}_p$ with positive valuation, the Laurent polynomial

$$d^2 \cdot a_1(t) \in \mathbb{Z}[t, t^{-1}]$$

has the same p-adic valuation when evaluated at ω and at α_0, with the exception of the primes occurring in α_0; we can take into account these as well and we obtain (possibly enlarging I_0)

$$\frac{1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_1(\alpha_0) - \alpha_1) - v_{\mathfrak{p}}(J_0)$$

for a positive integer J_0 depending on C only. We obtain

$$-v_{\mathfrak{p}}(x) \leq vv_{\mathfrak{p}}(a_1(\alpha_0) - \alpha_1) + vv_{\mathfrak{p}}(J_0)$$

from which we get, allowing $\mathfrak{p}$ to vary,

$$h_{fin}(x) \leq vh_{fin}((a_1(\alpha_0) - \alpha_1)^{-1}) + v \log J_0$$

which completes our estimates (even with an absolute bound on the height), unless the inequality above does not make sense: this happens precisely when $a_1(\alpha_0) = \alpha_1$.

We therefore assume $a_1(\alpha_0) = \alpha_1$ and we repeat the argument, with the series

$$y = \alpha_0 x^{u/v} + a_1(\alpha_0)x^{(u-1)/v} + \alpha_2 x^{(u-2)/v} + \dots$$

$$y = \omega x^{u/v} + a_1(\omega)x^{(u-1)/v} + a_2(\omega)x^{(u-2)/v} + \dots$$

$$uy - vxy' = a_1(\alpha_0)x^{(u-1)/v} + 2\alpha_2 x^{(u-2)/v} + 3\alpha_3 x^{(u-3)/v} + \dots$$

$$uy - vxy' = a_1(\omega)x^{(u-1)/v} + 2a_2(\omega)x^{(u-2)/v} + 3a_3(\omega)x^{(u-3)/v} + \dots$$

and we can compare the equations as above, recalling that

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha_0)) \geq -\frac{1}{v}v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(J_0)$$

and thus we obtain a similar estimate as above

$$\frac{2}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(\omega - \alpha_0) - v_{\mathfrak{p}}(G_1)$$

for an appropriate (fixed) positive integer G_1.
We now see how we are starting to apply a "bootstrap" argument: since

$$\omega - \alpha_0 \text{ divides } a_1(\omega) - a_1(\alpha_0)$$

in the sense that the ratio is a (fixed) polynomial in $\mathbb{Z}[t, t^{-1}, u, u^{-1}]$ evaluated at ω and α_0, divided then by d^2. This indeed implies the inequality

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha_0)) \geq -\frac{2}{v}v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(G_1')$$

for an appropriate fixed positive integer G_1'.
We now use this estimate together with the comparison of the equations for $uy - vxy'$ and we obtain instead

$$\frac{1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_2(\omega) - \alpha_2) - v_{\mathfrak{p}}(H_1)$$

for a constant H_1 depending on C only.
As above, we have that either

$$-v_{\mathfrak{p}}(x) \leq v_{\mathfrak{p}}(I_1)$$

for an appropriate positive integer I_1 or we can substitute α_0 instead of ω:

$$\frac{1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_2(\alpha_0) - \alpha_2) - v_{\mathfrak{p}}(J_1)$$

for a fixed positive integer J_1 depending on C only.
Again, as $\mathfrak{p}$ varies, we obtain

$$h_{fin}(x) \leq vh_{fin}((a_2(\alpha_0) - \alpha_2)^{-1}) + v \log J_1$$

which concludes, unless $a_2(\alpha_0) = \alpha_2$.

We can use this argument to show that $\alpha_i = a_i(\alpha_0)$ for every positive integer i. We show the next step.

We assume $a_1(\alpha_0) = \alpha_1$ and $a_2(\alpha_0) = \alpha_2$ and we have

$$y = \alpha_0 x^{u/v} + a_1(\alpha_0)x^{(u-1)/v} + a_2(\alpha_0)x^{(u-2)/v} + \dots$$

$$y = \omega x^{u/v} + a_1(\omega)x^{(u-1)/v} + a_2(\omega)x^{(u-2)/v} + \dots$$

$$uy - vxy' = a_1(\alpha_0)x^{(u-1)/v} + 2a_2(\alpha_0)x^{(u-2)/v} + 3\alpha_3 x^{(u-3)/v} + \dots$$

$$uy - vxy' = a_1(\omega)x^{(u-1)/v} + 2a_2(\omega)x^{(u-2)/v} + 3a_3(\omega)x^{(u-3)/v} + \dots$$

and we compare again the equations for y:

$$v_{\mathfrak{p}}(a_2(\omega) - a_2(\alpha_0)) \geq -\frac{1}{v} v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(J_1)$$

which, together with the previous bound

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha_0)) \geq -\frac{2}{v} v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(J_0)$$

implies

$$\frac{3}{v} v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(\omega - \alpha_0) - v_{\mathfrak{p}}(G_2)$$

for an appropriate (fixed) positive integer G_2.
Again,

$$\omega - \alpha_0 \text{ divides } a_2(\omega) - a_2(\alpha_0)$$

in the same sense as above and then

$$v_{\mathfrak{p}}(a_1(\omega) - a_1(\alpha_0)) \geq -\frac{3}{v} v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(G_2')$$

$$v_{\mathfrak{p}}(a_2(\omega) - a_2(\alpha_0)) \geq -\frac{3}{v} v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(G_2')$$

for an appropriate fixed positive integer G_2'.
We use again these estimates together with the comparison of the equations for $uy - vxy'$ and we obtain instead

$$\frac{1}{v} v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_3(\omega) - \alpha_3) - v_{\mathfrak{p}}(H_2)$$

for a constant H_2 depending on C only.
Again,

$$-v_{\mathfrak{p}}(x) \leq v_{\mathfrak{p}}(I_2)$$

for an appropriate positive integer I_2 or we can substitute α_0 instead of ω:

$$\frac{1}{v} v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_3(\alpha_0) - \alpha_3) - v_{\mathfrak{p}}(J_2)$$

for a fixed positive integer J_2 depending on C only.
Finally, as $\mathfrak{p}$ varies, we obtain

$$h_{fin}(x) \leq v h_{fin}((a_3(\alpha_0) - \alpha_3)^{-1}) + v \log J_2$$

which again implies $a_3(\alpha_0) = \alpha_3$.

This argument can be repeated indefinitely. Indeed, we obtain that, for every positive integer k, there exist positive integer constants G_k, G_k', H_k, I_k, J_k, depending on C only, such that

$$\text{if we suppose } a_i(\alpha_0) = \alpha_i \text{ for every } 1 \leq i \leq k$$

then

$$\frac{k+1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(\omega - \alpha_0) - v_{\mathfrak{p}}(G_k)$$

$$v_{\mathfrak{p}}(a_i(\omega) - a_i(\alpha_0)) \geq -\frac{k+1}{v}v_{\mathfrak{p}}(x) - v_{\mathfrak{p}}(G'_k) \text{ for every } 1 \leq i \leq k$$

$$\frac{1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_{k+1}(\omega) - \alpha_{k+1}) - v_{\mathfrak{p}}(H_k)$$

and finally

$$\frac{1}{v}v_{\mathfrak{p}}(x) \geq -v_{\mathfrak{p}}(a_{k+1}(\alpha_0) - \alpha_{k+1}) - v_{\mathfrak{p}}(J_k)$$

whenever

$$-v_{\mathfrak{p}}(x) \leq v_{\mathfrak{p}}(I_k)$$

which in particular implies our weakly bounded height bound unless $a_{k+1}(\alpha_0) = \alpha_{k+1}$.

4. We should now argue that if

$$a_i(\alpha_0) = a_i \text{ for every positive integer } i$$

then C is a modular curve. Indeed, let $f_s(x)$ be a holomorphic branch defined by the expansion (around ∞)

$$f_s(x) = sx^{u/v} + a_1(s)x^{(u-1)/v} + a_2(s)x^{(u-2)/v} + \dots$$

and we notice that for any

$$\omega = \exp(-2\pi i(q/r)) \text{ where } q \text{ and } r \text{ are coprime integers}$$

the function $f_\omega(x)$ is the branch of a modular curve corresponding to the linear transformation of the upper half-plane $\mathbb{H}$

$$\tau \to \frac{u}{v}\tau + \frac{q}{r}.$$

It seems very natural then to claim that $f_s(x)$ "corresponds" to the linear transformation

$$\tau \to \frac{u}{v}\tau - \frac{\log s}{2\pi i}$$

where we are fixing a determination of the logarithm; the correspondence is intended in the sense that

$$f_t(j(\tau)) = j\left(\frac{u}{v}\tau - \frac{\log s}{2\pi i}\right)$$

for every $\tau \in D$ with a sufficiently big imaginary part (where the v-th root, defining $x^{u/v}$, needs to be taken accordingly). This can be directly from the explicit construction of the Puiseux series for the modular curves we performed above, noting that the "q-expansion"

$$j\left(\frac{a}{d}\tau + \frac{b}{d}\right) = \frac{\omega}{q^{a/d}} + 744 + \frac{196884}{\omega}q^{a/d} + \dots$$

depends only on the relevant linear transformation.

We finally apply a functional transcendence result (namely, our usual Ax-Schanuel for the j-function, Theorem 3.4.9 by Pila-Tsimerman) that discards this possibility. If we let the holomorphic functions defined on an open subset of D

$$u_1(t) = t \text{ and } u_2(t) = \frac{u}{v}t - \frac{\log s}{2\pi i}$$

then the field

$$\mathbb{C}(t, u_1(t), u_2(t), j(u_1(t)), j(u_2(t)))$$

has transcendence degree (over $\mathbb{C}(t)$) at most 1, since

$$f_s(j(u_1(t))) = j(u_2(t)))$$

and $f_s(x)$ is by definition a branch of the algebraic curve C.
This indeed implies that C is special and our proof is completed: whenever C is not special the "bootstrap" procedure performed above cannot be continued indefinitely, hence for some positive integer i we have

$$a_i(\alpha_0) \neq \alpha_i$$

and thus we obtain an upper bound (implying weakly bounded height) for x.

Remark 6.3.3. *We point out that this very last argument is not effective: we have no criterion to find an i such that $a_i(\alpha_0) \neq \alpha_i$; we just know that, by functional transcendence, some must exist.*

We present a possible argument for the effectivity of the finite part of weakly bounded height; this is probably very out of reach, but it is connected to rather interesting arithmetic questions.

We refer to the simpler situation of the previous section. We could prove effectively that

$$v_{\mathfrak{p}}(\omega - \alpha_0) \geq -3v_{\mathfrak{p}}(x) \geq 3.$$

For the sake of simplicity let us assume α an integer; hence

$$p^3 \mid \alpha^n - 1.$$

We would wish to be able to obtain estimates on the biggest cube dividing $\alpha^n - 1$. There are two cases.

1. If

   $$p^3 \nmid \alpha^{p-1} - 1$$

   then $p \mid n$ and this immediately gives a bound on the height of the products of these admissible prime numbers (namely, the height is bounded by $\log n$) and similarly if we take into account multiplicities.

2. If

$$p^3 \mid \alpha^{p-1} - 1$$

the situation is much more complicated.
The prime numbers that satisfy the condition

$$p^2 \mid \alpha^{p-1} - 1$$

are called *Wieferich primes* if $\alpha = 2$ and they are expected to be rare; also, prime numbers of this kind appear in arithmetic versions (which are often harder) of unlikely intersections problems.
For instance, the expectation for only finitely many primes p to satisfy both

$$p^2 \mid 2^{p-1} - 1 \text{ and } p^2 \mid 3^{p-1} - 1$$

is an arithmetic version (i.e. over Spec $\mathbb{Z}$) of the results of section 2.3. We also refer to the similarity with the following Theorem of Masser-Zannier [43]:

considering the elliptic scheme, parameterised by λ, given by

$$\{(x, y, \lambda) \in \mathbb{A}^3 \mid y^2 = x(x-1)(x-\lambda)\}$$

the two sections defined by $x = 2$ and $x = 3$ both yield torsion points for only finitely many values of $\lambda \in \mathbb{A}^1$.

Our aim might be consequence of an appropriate Conjecture on the density of these Wieferich-type primes that might have an "arithmetic unlikely intersections" meaning.
For the primes p such that

$$p^3 \mid \alpha^{p-1} - 1$$

we would even expect finiteness, but we cannot say anything about these either.

Chapter 7

Further results

In this chapter we discuss some results related to "avoiding problems", although not falling in the previous realms.

7.1 Identically isogenous curves

In this section we discuss criteria to detect whether two curves are "identically isogenous".
We have been consistently concerned with the following condition: given two subvarieties $V, V' \subseteq S$ of a Shimura variety S, we need at least to require V and V' to be in some sense "unrelated" in order to hope to find points on V that avoid V'.
Clearly, it would be nice to have simple criteria to check whether V and V' are related or not.
We investigate this issue in the case of curves in $Y(1)^2$.

Given a curve $C \subseteq Y(1)^2$ and two positive integers m, n, we had defined already in section 3.4

$$C_{m,n} = \{(x, y) \in Y(1)^2 \mid \exists (x', y') \in C, \Phi_m(x, x') = 0, \Phi_n(y, y') = 0\}.$$

$C_{m,n}$ is the image of C through a correspondence of $Y(1)^2$ given by two modular polynomials.

Lemma 7.1.1. *$C_{m,n}$ is a closed curve.*

Proof. Let (x, y) be a point in the closure of $C_{m,n}$; we assume the contrapositive holds and we suppose $(x, y) \notin C_{m,n}$.
$C_{m,n}$ is the projection onto two coordinates of the closed curve C'

$$\{(x, y, x', y') \in Y(1)^4 \mid (x', y') \in C, \Phi_m(x, x') = 0, \Phi_n(y, y') = 0\}$$

and as (x, y) belongs to the (even euclidean) closure of $C_{m,n}$ there must be a sequence

$$(x_1, y_1, x_1', y_1'), (x_2, y_2, x_2', y_2'), \dots$$

of points of $C'(\mathbb{C})$ such that the following limits hold

$$\lim_{i \to +\infty} x_i = x$$

$$\lim_{i\to+\infty} y_i = y$$

and at least one of these two holds

$$\lim_{i\to+\infty} |x_i'| = +\infty;$$

$$\lim_{i\to+\infty} |y_i'| = +\infty.$$

If that was not the case, we could extract a subsequence converging to a point in C' whose image would be the very point (x, y).
This is not possible: indeed, for every bounded subset S of $\mathbb{C}$, there is a corresponding bounded subset $S' \subseteq \mathbb{C}$ such that whenever

$$\Phi_m(z_1, z_2) = 0 \text{ and } z_1 \in S \text{ then } z_2 \in S$$

and likewise for n. This can be deduced immediately by pulling back to the upper half-plane via j.
Since $C_{m,n}$ is an euclidean closed constructible set (in the sense of the language of $\mathbb{C}$, as a ring), it is also Zariski closed. This concludes the proof. □

Remark 7.1.2. *We point out that $C_{m,n}$ is not necessarily irreducible. Indeed, a strategy to construct a reducible $C_{m,n}$, that we also exploit later, is the following: let us fix an irreducible algebraic curve C and let $C' = C_{m,n}$; we observe that C is an irreducible component of $C'_{m,n}$ and as long as not both m and n are equal to 1 then $C'_{m,n}$ will be irreducible.*

Definition. *We say that two curves C and C' are isogenous if $C' \subseteq C_{m,n}$ for some positive integers m, n.*

Remark 7.1.3. *This is an equivalence relation as long as C and C' are meant to be irreducible.*

We recall that C and C' are isogenous if and only if the surface $C \times C' \subseteq Y(1)^2 \times Y(1)^2$ intersects a two-dimensional special subvariety of the shape

$$\{(x, y, x', y') \in Y(1)^4 \mid \Phi_m(x, x') = 0, \Phi_n(y, y') = 0\}$$

in a curve; this is, indeed, an unlikely intersection: we do not expect two "randomly chosen" curves to be isogenous.

We are mostly interested in the following two questions.

Question.

- ***A.*** *Given two curves $C, C' \subseteq Y(1)^2$, can we determine (effectively) whether they are isogenous?*
- ***B.*** *Given a "continuous family" $\mathcal{F}$ of curves in $Y(1)^2$ (e.g. all the curves of a given degree), can we find many curves which are not isogenous to any curve in $\mathcal{F}$?*

This kind of questions arose in the first instances of the André-Oort Conjecture; we refer to the paper by André [1] for the first of such results. Edixhoven, in [20], proved that if $C \subseteq Y(1)^2$ is not special, then for any sufficiently big prime p we have $C \nsubseteq C_{p,p}$. Edixhoven's reasoning does not seem immediately adaptable to our case, since his proof uses a number of group theoretic properties that become convoluted whenever m and n have small prime factors.
Here we present three different approaches. First, we present arguments using o-minimality, which are very neat, but have the big disadvantage of being ineffective. Second, we present a direct method exploiting the points at infinity, which is easy to check but has a rather narrow scope of application. Third, we present the most powerful approach, which settles the issue completely and effectively through the result already aforementioned by Binyamini-Daw, in [9].

Let us deal with the two questions. Let $C \subseteq Y(1)^2$ be a curve and let $\mathcal{F}$ be a "continuous family" of curves of $Y(1)^2$ with fixed "algebro-geometric" degree.
We show below that there are only finitely many curves in $\mathcal{F}$ which are isogenous to C, unless C is either a "vertical" or "horizontal" line.

1. Here we use o-minimality arguments.
   First, if C is special then the result is obvious since a special curve can be isogenous only to another special curve; this is an immediate consequence of the fact that every non-special curve contains a point (x, y) (possibly not defined over $\overline{\mathbb{Q}}$) whose two coordinates are not isogenous. As the geometric degree (in the sense of section 3.7) of the modular curve Y_n is $2\psi(n)$, only finitely many special curves can be represented by some point in $\mathcal{F}$.
   We can therefore assume C non-special. We argue defining the following set S, where we denote with C_p the curve represented by the point $p \in \mathcal{F}$:

$$\{(a,b,c,d,a',b',c',d') \in \mathbb{R}^8, p \in \mathcal{F} \mid$$

$$\exists \text{ "small" open subsets } U \subseteq C(\mathbb{C}) \text{ and } V \subseteq C_p(\mathbb{C})$$

$$\text{such that } \forall (x,y) \in U \ \exists (x',y') \in V$$

$$\frac{aj^{-1}(x)+b}{cj^{-1}(x)+d} = j^{-1}(x') \text{ and } \frac{a'j^{-1}(y)+b'}{c'j^{-1}(y)+d'} = j^{-1}(y')\}$$

   where, as usual, j is restricted to the standard fundamental domain $D \subseteq \mathbb{H}$.
   The key observation here is that if $((a,b,c,d,a',b',c',d'),p) \in S$, then a,b,c,d and a',b',c',d' must be integers (up to rescaling, in the sense that $(\lambda a, \lambda b, \lambda c, \lambda d)$ acts on $\mathbb{H}$ in the same way as (a,b,c,d) for any nonzero scalar λ), as a consequence of functional transcendence. Indeed, we can construct algebraic functions $x(t), y(t)$ and holomorphic functions $x'(t), y'(t)$ defined on U satisfying

$$\frac{aj^{-1}(x(t))+b}{cj^{-1}(x(t))+d} = j^{-1}(x'(t)) \text{ and } \frac{a'j^{-1}(y(t))+b'}{c'j^{-1}(y(t))+d'} = j^{-1}(y'(t)).$$

The Ax-Schanuel Theorem for the j-function (Theorem 3.4.9) by Pila-Tsimerman where, as usual, we set

$$u(t) = j^{-1}(x(t)), v(t) = j^{-1}(y(t)), u'(t) = j^{-1}(x'(t)), v'(t) = j^{-1}(y'(t))$$

and we ignore the problems at the boundary of D (that would disappear choosing a slightly bigger open enlargement of D), implies that the field

$$\mathbb{C}(t, u(t), v(t), u'(t), v'(t), x(t), y(t), x'(t), y'(t))$$

has transcendence degree over $\mathbb{C}(t)$ at least 4 unless something "weakly special" happens. In this case, the relevant transcendence degree is at most 3, since $x(t)$ and $y(t)$ are algebraic, $u(t)$ and $u'(t)$ (and likewise $v(t)$ and $v'(t)$) are related by an algebraic equation by definition of S, while $x'(t)$ and $y'(t)$ parameterise points on the curve C_p.
Thus, something "weakly special" must be occurring. As we briefly observed above, a constant coordinate, i.e. C being a "vertical" or "horizonatal" line is indeed an exception. The cases $x(t)$ and $y(t)$ (or $x'(t)$ and $y'(t)$) identically related by a modular polynomial fall into the realm of C special of C_p special (which in turn implies that C is special, as we have argued before), that we have treated already.
Thus, we can assume $x(t)$ and $x'(t)$ (rather than $y(t)$ and $y'(t)$, without loss of generality) related by a modular polynomial:

$$\Phi_m(x(t), x'(t)) = 0 \text{ for every } t \in U$$

for some positive integer m. This implies that a, b, c, d need to be (up to rescaling) integers, since $x(t)$ is nonconstant. Indeed, if

$$\frac{a\tau + b}{c\tau + d} = \frac{A\tau + B}{C\tau + D}$$

then

$$\tau^2(aC - cA) + \tau(aD + bC - cB - dA) + (bD - dB) = 0$$

and, if τ is allowed to vary among infinitely many values as above, then (A, B, C, D) must be equivalent to (a, b, c, d) up to rescaling; we are implicitly using that $ad - bc \neq 0$. Thus, only one value of (a, b, c, d) (up to rescaling) is admissible and it is indeed given by the vanishing of the modular polynomial as above.
We apply again Ax-Schanuel for the j-function, pointing out that by the integrality of a, b, c, d then $x'(t)$ and hence also $y'(t)$ are algebraic functions of t, hence the field

$$\mathbb{C}(t, v(t), v'(t), y(t), y'(t))$$

has transcendence degree at most 1 over $\mathbb{C}(t)$ and therefore, by the exact same reasoning as above, the entries a', b', c', d' are integers (up to rescaling).

Finally S is a definable set such that every point $((a, b, c, d, a', b', c', d'), p) \in S$ necessarily

satisfies (a, b, c, d) and (a', b', c', d') integer vectors (up to rescaling). By o-minimality, then (a, b, c, d) and (a', b', c', d') can only assume finitely many values (again, up to rescaling); indeed, any countable definable set is finite.
Thus, there is an absolute upper bound on any m, n such that $C_{m,n} \in \mathcal{F}$ and, in turn, there are only finitely many $p \in \mathcal{F}$ such that C is isogenous to C_p.

This proof settles question B., but we are not able to obtain any information on A., as our methods are not effective.

Remark 7.1.4. *We can even obtain more: suppose that C is not fixed, but allowed to vary in a "continuous family" $\mathcal{F}'$ (of curves of bounded degree); we aim to find some $C \in \mathcal{F}'$ which is not isogenous to any curve parameterised by a point of $\mathcal{F}$.*
The definable set S described in the proof above can be defined uniformly along all the family $\mathcal{F}'$ and hence there is an upper bound on the degrees m, n of the connecting isogenies between any element of $\mathcal{F}$ and any element of $\mathcal{F}'$.
Therefore, in order to produce a curve in $\mathcal{F}'$ which is not isogenous to any curve in $\mathcal{F}$ we just need to find a "bigger" family $\mathcal{F}'$ (meaning that the dimension of the Chow variety corresponding to $\mathcal{F}'$ is bigger than that corresponding to $\mathcal{F}$). This condition is indeed necessary: for instance, the case $\mathcal{F}' = \mathcal{F}$ is unavoidable.
We can also choose our "non-isogenous curve" in $\mathcal{F}'$ to be defined over a "small" field (namely, the field of definition of $\mathcal{F}'$).

2. We have a good understanding of the action of isogenies on the points at infinity if we embed $Y(1)^2 \subseteq (\mathbb{P}_1)^2$ (in the usual way), thanks to the following observation that we already made during the proof of Lemma 7.1.1: if z_1 and z_2 are related by the modular equation
$$\Phi_n(z_1, z_2) = 0$$
for a fixed positive integer n, then $|z_1|$ tends to $+\infty$ if and only if $|z_2|$ does.
From this we can deduce that if C and C' are isogenous curves in $Y(1)^2 \subseteq (\mathbb{P}_1)^2$, then:

   - (∞, ∞) is a point at infinity of C if and only if it is a point at infinity of C';
   - (a, ∞) is a point at infinity of C if and only if there is some b, isogenous to a, such that (b, ∞) is a point at infinity of C';
   - analogously for the symmetric situation (∞, a)...

   For instance, the two curves $\{(x, y) \mid xy = 1\}$ and $\{(x, y) \mid y = x+1\}$ cannot be isogenous, since the first has points at infinity $(\infty, 0)$ and $(0, \infty)$, while the second only has (∞, ∞). This criterion is handy in some situation when it can be used to answer questions A. and B., but its scope of application is definitely limited.

3. We can settle the question completely using again the result of Binyamini-Daw of [9], which bounds the "algebro-geometric" degree of the "weakly optimal" subvarieties of any variety in $Y(1)^n$.
   We suppose that the curve E, with no constant coordinate, satisfies

$$E \subseteq C \times C' \cap \{(x, y, x', y') \mid \Phi_m(x, x') = 0, \Phi_n(y, y') = 0\}$$

   and, by the Binyamini-Daw result, the "algebro-geometric" degree of E is bounded above in terms of the geometric degrees of C and C' only.
   We can construct nonconstant morphisms, given by coordinate projections, from E to

$$\{(x, x') \mid \Phi_m(x, x') = 0\}$$

   and from the latter to $Y(1)$; this last morphism has degree $\psi(m)$. Since the degree of the composition of these two morphisms (which is just the intersection number of E intersected with an appropriate hyperplane) is bounded by the Binyamini-Daw result, also m is bounded; the same holds for n.
   Thus, we obtain an effective upper bound on m and n depending on $\mathcal{F}$ and $\mathcal{F}'$ only, where $\mathcal{F}'$ is a "continuous family" from which we can pick C. Effective answers to questions A. and B. follow.

It is interesting in itself to understand which curves C are isogenous (in a nontrivial manner) to themselves, i.e. there exist m, n, not both 1, such that $C \subseteq C_{m,n}$.

Remark 7.1.5. *It follows from the o-minimality argument in 1. above that there are only finitely many pairs (m, n) such that $C \subseteq C_{m,n}$ whenever C is not weakly special: setting $\mathcal{F} = \{C\}$ in the proof, then the pairs of admissible integers (m, n) are seen to be a definable subset of $\mathbb{R}^2$, which is therefore finite.*

Weakly special curves are indeed "self-isogenous", but there are other examples. Let n be a positive integer and let C be defined as

$$\{(x, y) \mid \Phi_n(x, y^2) = 0\}.$$

Then $C \subseteq C_{n^2,1}$. Indeed, if $(x, y) \in C$, then there is some x' such that

$$\Phi_{n^2}(x, x') = 0$$

and $(x', y) \in C$.

For a while we considered possible that all of these "counterexamples" required either m or n equal to 1; Tsimerman kindly explained us how to construct examples where m and n are arbitrary squares.
Let us take a "generic" curve C, in the sense that $C_{m,n}$ is irreducible: this is "generally" true by an application of Bertini's Theorem.

If we let $E = C_{m,n}$, we claim that $E \subseteq E_{m^2,n^2}$.
Let us take any point $(p, q) \in E$; we need to show that there is a point $(p', q') \in E$ such that

$$\Phi_{m^2}(p, p') = 0 \text{ and } \Phi_{n^2}(q, q') = 0.$$

By definition of E there is a point $(p'', q'') \in C$ such that

$$\Phi_m(p, p'') = 0 \text{ and } \Phi_n(q, q'') = 0$$

and we can choose p' and q' so that

$$\Phi_m(p'', p') = 0 \text{ and } \Phi_n(q'', q') = 0;$$

$$\Phi_{m^2}(p, p') = 0 \text{ and } \Phi_{n^2}(q, q') = 0.$$

Therefore, $(p', q') \in E$ by definition.

Remark 7.1.6. *This argument also implies that if C and C' are isogenous, then they are either isomorphic or there are positive integers m, n, not both 1, such that $C \subseteq C_{m,n}$ or $C' \subseteq C'_{m,n}$.*

We could ask whether these conditions on m and n are indeed necessary.

Question. *If $C \subseteq Y(1)^2$ is such that $C \subseteq C_{m,n}$, then:*

1. *Do m, n have to be squares?*
2. *Does mn have to be a square?*

We do not strongly believe 1. to be true, but we do not have counterexamples. We give a partial answer to 2. with a rather strong condition on the points at infinity of our curve.

Proposition 7.1.7. *Let $C \subseteq Y(1)^2$ be a curve over $\overline{\mathbb{Q}}$ and suppose that the ratios x/y and y/x are bounded outside of a compact (for the euclidean topology) subset of $C(\mathbb{C})$. If m and n are positive integers such that $C \subseteq C_{m,n}$, then mn is a square.*

Remark 7.1.8. *The condition requiring the ratios x/y and y/x to be bounded outside of an euclidean compact set is the same as requiring each Puiseux series expansions of C at infinity to be of the shape*

$$y = cx + O(x)$$

for a nonzero $c \in \overline{\mathbb{Q}}$ (possibly with a different c for each branch).
Such condition can also be expressed in terms of the points at infinity of $C \subseteq Y(1)^2 \subseteq \mathbb{P}_2$; we prefer not to insist on this, since usually it is more natural to consider the inclusion $Y(1)^2 \subseteq (\mathbb{P}_1)^2$.

Proof. This is a consequence of the Proposition 7.1.9 below.
Let $(x, y) \in C(\overline{\mathbb{Q}})$ which is isogenous to some $(x', y') \in C(\overline{\mathbb{Q}})$ via

$$\Phi_m(x, x') = 0 \text{ and } \Phi_n(y, y') = 0$$

and let $\mathfrak{p}$ be a prime of a number field containing $\mathbb{Q}(x, y, x', y')$.
We can assume $\mathfrak{p}$ not to occur into the prime factorisation of any c which appears as a leading term of any Puiseux series expansion of C (as we describe in the Remark above) and moreover not to occur with a negative exponent in the prime factorisation of any coefficient of any Puiseux series expansion of C; we point out to the already quoted observation by Eisenstein in [57] that guarantees that this is possible.
We can finally choose x so that $v_{\mathfrak{p}}(x) < 0$ and, by our condition on the Puiseux series, we have

$$v_{\mathfrak{p}}(x) = v_{\mathfrak{p}}(y)$$

and, by Proposition 7.1.9, both

$$mv_{\mathfrak{p}}(x)/v_{\mathfrak{p}}(x')$$

$$nv_{\mathfrak{p}}(y)/v_{\mathfrak{p}}(y')$$

are squares; thus, again by our constraints on the Puiseux series, we have

$$v_{\mathfrak{p}}(x') = v_{\mathfrak{p}}(y')$$

and therefore mn needs to be a square. □

Proposition 7.1.9. *Let m be a positive integer, let $x, x' \in \overline{\mathbb{Q}}$ be algebraic numbers and let $\mathfrak{p}$ be a prime ideal of a number field containing $\mathbb{Q}(x, x')$. Suppose moreover*

$$\Phi_m(x, x') = 0 \text{ and } v_{\mathfrak{p}}(x) < 0.$$

Then

$$v_{\mathfrak{p}}(x') < 0 \text{ and } m \cdot v_{\mathfrak{p}}(x)/v_{\mathfrak{p}}(x') \text{ is the square of a nonzero rational number.}$$

Proof. The proof follows by looking explicitly at the coefficients of the modular polynomials, namely at their "upper-right convex hull" in a fashion similar to the method of Newton polygons.

We start with an example, where we showcase the coefficients of the modular polynomial $\Phi_{12}(x, y)$ forgetting about the monomials $x^a y^b$ such that there exist positive integers a_1, b_1 and a_2, b_2 with a real $0 \leq \lambda \leq 1$ satisfying:

- $a < \lambda a_1 + (1 - \lambda)a_2$ and $b < \lambda b_1 + (1 - \lambda)b_2$;
- the coefficients of $x^{a_1} y^{b_1}$ and of $x^{a_2} y^{b_2}$ in the expansion of $\Phi_{12}(x, y)$ is nonzero.

These monomials (which, we recall, have integer coefficients) are labelled with a question mark below. The remaining monomials are the "upper-right convex hull".

$$\begin{pmatrix} x^{24} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ ? & ? & -x^{23}y^{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ ? & ? & ? & ? & -x^{22}y^{15} & 0 & 0 & 0 & 0 \\ ? & ? & ? & ? & ? & x^{19}y^{19} & 0 & 0 & 0 \\ ? & ? & ? & ? & ? & ? & -x^{15}y^{22} & 0 & 0 \\ ? & ? & ? & ? & ? & ? & ? & 0 & 0 \\ ? & ? & ? & ? & ? & ? & ? & -x^{12}y^{23} & 0 \\ ? & ? & ? & ? & ? & ? & ? & ? & 0 \\ ? & ? & ? & ? & ? & ? & ? & ? & y^{24} \end{pmatrix}.$$

We give the following interpretation of this string of coefficients. If we let either one among (or, equivalently, both) x and y tend to infinity while satisfying $\Phi_{12}(x,y) = 0$, then one of the following must hold (where, as usual, we are restricting j to the standard fundamental domain D):

- $$\begin{pmatrix} 1 & b \\ 0 & 12 \end{pmatrix} j^{-1}(x) = j^{-1}(y)$$

  for any integer $0 \leq b < 12$, which correspond to the two monomials $x^{24} - x^{23}y^{12}$, in the sense that $x - y^{12}$ will be "relatively small";

- $$\begin{pmatrix} 2 & b \\ 0 & 6 \end{pmatrix} j^{-1}(x) = j^{-1}(y)$$

  for any *odd* integer $0 \leq b < 6$, which correspond to the two monomials $-x^{23}y^{12} - x^{22}y^{15}$, in the sense that $x + y^3$ will be "relatively small";

- $$\begin{pmatrix} 3 & b \\ 0 & 4 \end{pmatrix} j^{-1}(x) = j^{-1}(y)$$

  for any integer $0 \leq b < 4$, which correspond to the two monomials $-x^{22}y^{15} + x^{19}y^{19}$, in the sense that $x^3 - y^4$ will be "relatively small";

- $$\begin{pmatrix} 4 & b \\ 0 & 3 \end{pmatrix} j^{-1}(x) = j^{-1}(y)$$

  for any integer $0 \leq b < 3$, which correspond to the two monomials $x^{19}y^{19} - x^{15}y^{22}$, in the sense that $x^4 - y^3$ will be "relatively small";

- $$\begin{pmatrix} 6 & 1 \\ 0 & 2 \end{pmatrix} j^{-1}(x) = j^{-1}(y)$$

  which correspond to the two monomials $-x^{15}y^{22} - x^{12}y^{23}$, in the sense that $x^3 + y$ will be "relatively small";

- $$\begin{pmatrix} 12 & 0 \\ 0 & 1 \end{pmatrix} j^{-1}(x) = j^{-1}(y)$$

  which correspond to the two monomials $-x^{12}y^{23} + y^{24}$, in the sense that $x^{12} - y$ will be "relatively small".

The notion of "relatively small" can be made precise in a p-adic context: suppose that $x, y \in \mathbb{C}_p$ and that, satisfying $\Phi_{12}(x, y) = 0$, either one (or, equivalently, both) of them have a negative p-adic valuation. Then

$$v_p(x)/v_p(y) \in \{1/12, 1/3, 3/4, 4/3, 3, 12\}$$

otherwise (by integrality of the coefficients of the monomials marked with a question mark) exactly one among the monomials

$$x^{24}, -x^{23}y^{12}, -x^{22}y^{15}, x^{19}y^{19}, -x^{15}y^{22}, -x^{12}y^{23}, y^{24}$$

would have minimal p-adic valuation, contradicting $\Phi_{12}(x, y) = 0$.
Moreover, we obtain some "cancellation". Let us take for instance $v_p(x) = -4$ and $v_p(y) = -3$. We directly observe that every monomial M occurring in the expansion of $\Phi_{12}(x, y)$ satisfies $v_p(M) > -133$, with the only two exceptions

$$v_p(-x^{22}y^{15}) = -133 \text{ and } v_p(x^{19}y^{19}) = -133$$

and, in particular, by $\Phi_{12}(x, y) = 0$,

$$v_p(-x^{22}y^{15} + x^{19}y^{19}) > -133$$

i.e. there is some cancellation occurring in $x^3 - y^4$ (and in this sense it turns to be "relatively small").

We now make our claim on the structure of this "upper-right convex hull" precise. Let m be a positive integer. The coefficients appearing in the "upper-right convex hull" of the expansion of $\Phi_m(x, y)$ are divided into "strings"; each "string" corresponds to a choice of positive integers a, d such that $ad = m$ and

a monomial $c \cdot x^p y^q$ occurring in the expansion of $\Phi_m(x, y)$ appears in the (a, d)-string if and only if $aq + dp$ is maximal among all the monomials.

Referring to the example above, the $(2, 6)$-string consists of the polynomial $-x^{23}y^{12} - x^{22}y^{15}$.

Each choice of a, d satisfying $ad = m$ corresponds to all the matrices

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

for any b satisfying $\gcd(a,b,d)=1$ and $0 \leq b < d$.
Investigating the action of these matrices on the upper half-plane we observe that, if we perform a first order approximation of the j-function as

$$j(\tau) = \exp(-2\pi i\tau)$$

for the q-expansion then $j^{-1}(x)$ and $j^{-1}(y)$ being related by such a matrix implies

$$\Psi_g(x^{a'}, y^{d'}) = 0$$

where $g = \gcd(a,d)$ and $a = g \cdot a', d = g \cdot d'$; the polynomial $\Psi_g(\cdot,\cdot)$ is the usual cyclotomic polynomial (in the sense of $\mathbb{G}_{\mathrm{m}}$).

Summarising, we claim that the coefficients of the "upper-right convex hull" are exactly those that arise in the (a,d)-strings, whenever $ad = m$; each (a,d)-string corresponds to a polynomial

$$x^p y^q \Psi_g(x^{a'}, y^{d'})$$

where p and q are suitable natural numbers, $g = \gcd(a,d)$ and $a = g \cdot a', d = g \cdot d'$.

We make a quick example before giving arguments supporting our claim. Below we show the "upper-right convex hull" expansion of $\Phi_9(x,y)$

$$\begin{pmatrix}
x^{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
? & ? & ? & ? & ? & ? & -x^{11}y^9 & 0 & 0 & 0 \\
? & ? & ? & ? & ? & ? & ? & -x^{10}y^{10} & 0 & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & -x^9y^{11} & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & ? & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & ? & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & ? & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & ? & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & ? & 0 \\
? & ? & ? & ? & ? & ? & ? & ? & ? & y^{12}
\end{pmatrix}$$

and we observe that the $(3,3)$-string consists indeed of

$$-x^{11}y^9 - x^{10}y^{10} - x^9y^{11} = -x^9y^9 \cdot \Psi_3(x,y)$$

which takes into account three coefficients, rather than two as it always happened for $m = 12$.

The proof of the claim is not conceptually complicated. On one hand, we observe that the Puiseux series expansion (in a power of y rather than in y) of the modular curve defined as the zero locus of $\Phi_m(x,y) = 0$ corresponding to the matrix (acting on D)

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

where $ad = m$, $\gcd(b,g) = 1$, with $g = \gcd(a,d)$ and $a = g \cdot a', d = g \cdot d'$, starts as

$$y^{d'} = \omega \cdot x^{a'} + \ldots$$

where, as b varies, ω can be taken as any primitive g-th rooth of unity.
Using the usual properties of integrality of the coefficients of Puiseux series of algebraic functions, as we have observed previously, choosing x with negative valuation at a sufficiently big prime shows that the polynomial

$$\Psi_g(x^{a'}, y^{d'})$$

occurs as a factor in the polynomial given by some (A, D)-string (in principle even with $(A, D) \neq (a, d)$, although it is not hard to see that they are equal).

This quick observation allows almost to conclude. In order for everything to add up together, the only step missing now is proving that

$$\sum_{ad=m} \varphi(g)a' = \psi(m)$$

and this would automatically force the sequence of the (a, d)-strings to constitute the whole "upper-right convex hull".

We rewrite

$$\sum_{ad=m} \varphi(g)a' = \sum_{a|m} \frac{\varphi(g)}{g} a = \sum_{a|m} \sum_{f|a, \frac{m}{a}} \mu(f) \cdot \frac{a}{g} \cdot \frac{g}{f} = \sum_{a|m} \sum_{f|a, \frac{m}{a}} a \cdot \frac{\mu(f)}{f} =$$

$$= \sum_{f^2|m} \sum_{x|\frac{m}{f^2}} xf \cdot \frac{\mu(f)}{f} = \sum_{f^2|m} \sum_{x|\frac{m}{f^2}} x\mu(f) = \sum_{x|m} \sum_{f^2|\frac{m}{x}} x\mu(f) = \sum_{\substack{x|m \\ \frac{m}{x} \text{ is squarefree}}} x = \psi(m)$$

where we are using that

$$\sum_{e|n} \mu(e) = 0$$

whenever $n \geq 2$. This concludes the proof.

□

7.2 Avoiding problems in abelian varieties

In this section we briefly investigate "avoiding problems" in the context of a single abelian variety.
A prototypical "avoiding problem" is the following.

Question. *Let A be an abelian surface over $\overline{\mathbb{Q}}$ and let $C_1, C_2 \subseteq A$ be curves over $\overline{\mathbb{Q}}$ which are "unrelated" (in the sense that for no nonzero integers m, n the intersection $[m]C_1 \cap [n]C_2$ is positive-dimensional).*
Can we prove that "most" $P \in C_1(\overline{\mathbb{Q}})$ satisfy the following: there are no nonzero integers m, n and a point $Q \in C_2(\overline{\mathbb{Q}})$ with

$$[m]P = [n]Q?$$

This question admits a positive answer by a Theorem of Habegger in [26], which roughly states that the intersection of a subvariety of an abelian variety with special varieties of complementary dimension is a set of bounded height, under an appropriate "unrelatedness" condition. In this particular case, we are intersecting the surface $C_1 \times C_2$ with the two-dimensional special subvariety defined by

$$[m](\cdot) = [n](\cdot).$$

The result of Habegger is a natural translation of a result of Bombieri-Masser-Zannier for powers of $\mathbb{G}_{\mathrm{m}}$ (we point out to [10] and its generalisation by Habegger [27]).

The result of Habegger pretty much settles all the problems of the following shape:

let A be an abelian variety over $\overline{\mathbb{Q}}$ and let $V_1, V_2 \subseteq A$ be two "unrelated" varieties. Can we find a point on $V_1(\overline{\mathbb{Q}})$ which is "unrelated" to any point of $V_2(\overline{\mathbb{Q}})$?

The reasoning is roughly the following: we either shrink V_1 or expand V_2 until the dimensional requirements are obtained for the Theorem of Habegger to hold.

We point out that, in contrast to the Question above, proving such a property "most" points of $V_1(\overline{\mathbb{Q}})$ is not an immediate consequence of the Theorem of Habegger; indeed, we are not allowed to shrink V_1.
We can prove a partial result of this kind, where we only consider "backwards orbits".

Proposition 7.2.1. *Let A be an abelian variety over $\overline{\mathbb{Q}}$ and let $V_1, V_2 \subseteq A$ be two "unrelated" varieties over $\overline{\mathbb{Q}}$. Then, for d sufficiently big, "most" points P of $V_1(\overline{\mathbb{Q}})$ (of degree $\leq d$) satisfy the following property: for every positive integer n*

$$[n]^{-1}P \cap V_2 = \varnothing.$$

This Proposition is based on the following result of ours, in [6], which we quote as Theorem 7.2.2 below; this is an abelian analogue of a modular Fermat equation that was studied by Pila in [50].

We say that a point $P \in A(\overline{\mathbb{Q}})$ is "primitive" if its field of definition is K and P is indivisible in the Mordell-Weil group $A(K)$.
We point out that "most" points of $A(\overline{\mathbb{Q}})$ (of degree less than d, for d sufficiently big) are primitive: the number of points with height less than h will be roughly $\exp(ch)$ and the points of the form "2 times a point defined over the same field" will be roughly $\exp(ch/4)\ldots$
We have the following.

Theorem 7.2.2. *Let A be an abelian variety of dimension g, defined over the number field K. There is a positive constant c, depending effectively on A and $[K : \mathbb{Q}]$, such that, for every primitive point $P \in A(\overline{\mathbb{Q}})$ with $[K(P) : K] \leq d$ and every point $Q \in A(\overline{\mathbb{Q}})$, if*

$$[n]Q = P$$

for some positive integer $n \geq 2$, then

$$[K(Q) : K] \geq c \cdot \sqrt{d} \cdot n^{1/(4g)} / \sqrt{\log n}.$$

Remark 7.2.3. *We point out that the bound above does not depend on the height of P.*

This is a consequence of the estimates of David and Masser, who give lower bounds on the degree of the field of definition of a torsion point in an abelian variety in terms of its order (we refer to the works [39] and [16]); we also need a combinatorial argument.
We skip the full proof, only giving an example when $n = 2p$, where p is an odd prime number. Let us consider Q', a Galois conjugate over K of Q. The point $Q' - Q$ is torsion of order equal to some m dividing n, because

$$[n]Q = [n]Q' = P$$

and we assume such order to be maximal among the choices of the conjugate Q'.

- If the order m is either p or $2p$, then we estimate the degree of the field of definition of $Q - Q'$ using the results by Masser or David. We obtain a lower bound of the size

  $$[K(Q' - Q) : K] \geq n^{1/(2g) - o(1)}$$

  and hence we can bound from below the degree of the field of definition of Q (and Q') over K by a quantity comparable to

  $$[K(Q) : K] \geq n^{1/(4g) - o(1)}.$$

- The order m cannot be 1 since P is primitive and, if $m = 2$, then

  $$[2]Q = [2]Q' = [2]Q''$$

  for any conjugate Q'' of Q over k by our assumption of maximality on the order. Hence, $[p]([2]Q) = P$, which contradicts P being primitive.

The proof for a general n follows from the following combinatorial statement.

Proposition 7.2.4. *Let G be a finite abelian group and let N be the minimum positive integer such that $N \cdot g = 0$ for every $g \in G$. Let S be a subset of G with the following properties.*

1. *The lcm of the orders of the elements of S is N.*
2. *For every $a, b \in S$, there exists $c \in S$ such that $a - b$ and c have the same order.*

Then, S contains an element of order $\geq \sqrt{N}$.

We now briefly sketch how Proposition 7.2.1 follows from Theorem 7.2.2.

- We can assume P to be primitive (over its field of definition) since "most" points of $V_1(\overline{\mathbb{Q}})$ satisfy this property.

- Theorem 7.2.2 implies that each element Q of $[n]^{-1}P$ has a field of definition with "big" degree in terms of n.
- We can therefore apply the Pila-Wilkie Theorem 2.2.5 taking into account all the Galois conjugates of Q; we need to treat carefully the maximal weakly special subvarieties contained in V_2 (similarly to what we did during the proof of Theorem 3.6.7).
- After obtaining an upper bound on n, we can finally use the hypothesis of V_1 and V_2 being unrelated in order to prove that the "isogenous images" $[n]V_2(\overline{\mathbb{Q}})$ are "far from covering" $V_1(\overline{\mathbb{Q}})$.

Remark 7.2.5. *Theorem 7.2.2 can also be used to give partial answers to the question above for C_1 and C_2 which are not "unrelated", for instance if $C_1 = C_2$. The result of Habegger does not cover this case; we refer to the paper of Bays-Habegger* [8]*, in the context of $\mathbb{G}_{\mathrm{m}}$, where a similar problem is treated.*

In the abelian context there is also the possibility to formulate "relative problems", but here much less seems to be known.
For instance, we can consider the family of curves parameterised by t

$$C_t : \{(x, y) \mid y^2 = x^6 + x + t\}$$

and their jacobians J_t. We consider the two points at infinity ∞^+ and ∞^- (corresponding to the points at infinity of $y^2 = x^6$, as can we "forget" terms of smaller degree) and this allows us to produce the point

$$\delta_t = (\infty^+) - (\infty^-)$$

on each C_t, together with an embedding in J_t

$$P \to (P) - (\infty^-).$$

The following question was raised in a paper of Masser-Zannier, [44].

Question. *Can we find values of $t \in \overline{\mathbb{Q}}$ for which there is no $n \geq 2$ such that $[n]\delta_t \in C_t$?*

Such values were found by Flynn in the same paper using Chabauty-Kim methods; these methods are powerful but do not seem to provide a general answer to this kind of problems.

This question is also related to continued fractions in polynomials; namely, the question is equivalent to finding some t for which all the partial quotients of the continued fraction of $\sqrt{x^6 + x + t}$ have degree 1.
For more on the beautiful connection between polynomial continued fractions and Diophantine Geometry, we point to the paper by Zannier, [66], or the notes by Capuano-Jossen-Karolus-Veneziano, [11].

7.3 Avoiding problems and Hecke orbits

In this section we discuss a rather different kind of "avoiding problems", related to a fixed Hecke orbit.

Definition. *The Hecke orbit of an algebraic number $\alpha \in \overline{\mathbb{Q}}$ is the set*

$$\{\beta \in \overline{\mathbb{Q}} \mid \exists m \text{ such that } \Phi_m(\alpha, \beta) = 0\}.$$

We may ask the following question.

Question.
- ***A.*** *Does there exist an algebraic number $\alpha \in \overline{\mathbb{Q}}$ such that for no β in its Hecke orbit we have $h(\beta) < 1000$?*
- ***B.*** *Does there exist an algebraic integer $\alpha \in \overline{\mathbb{Q}}$ such that no β in its Hecke is a unit?*

Remark 7.3.1. *There is a "special" choice for α that would answer to question A.: a carefully chosen singular modulus would work, since we have a good control over the height of singular moduli.*
We later show that we can obtain a much more "general" α.

We have an affirmative answer to the two questions.

Proposition 7.3.2. *For "most" integer numbers $q \in \mathbb{Q}$, there is no p in the same Hecke orbit as q such that:*

- *$h(p) < 1000$;*
- *p is a unit.*

We only give a sketch of the proof.

- We use a result of Zywina (which is a uniform version of Serre's Open Image Theorem), [68], which implies that there is a positive constant c such that, for "most" choices of q, the action of the absolute Galois group of $\mathbb{Q}$ over any Tate module attached to E_q (any elliptic curve with j-invariant equal to q) has image which is a subgroup of GL_2 of index at most c. In particular,

$$\text{whenever } \Phi_m(p, q) = 0 \text{ then } [\mathbb{Q}(p) : \mathbb{Q}] \geq \psi(m)/C.$$

- Thanks to an equidistribution result by Clozel-Ullmo, in [13] (or explained in the notes by Ullmo [64]), the sets

$$\{p \in D \mid \Phi_m(p, q) = 0\}$$

tend to the hyperbolic measure on D, as m tends to $+\infty$.
By the previous discussion, the same equidistribution occurs also for the Galois orbits.

- We can already give lower bounds on the height, as we have enough "big" (which in D is the same as being "high") points.
  In order to give the correct lower bound on the norms we need to ensure that p is sufficiently away from 0, i.e. $j^{-1}(p)$ needs to be away from the "corners" of D, which are roots of unity (of order either 3 or 6). This can be sorted using Linear Forms in elliptic Logarithms; we refer to [15], or the form used by Habegger in [28] for the same purpose as ours.

Remark 7.3.3. *We point out that a largely similar result, with a partly identical proof, has been obtained already by Schmid, in* [60]*, who proved, among other things, that any Hecke orbit contains only finitely many units.*
Our result has the additional feature of appealing to Zywina's Theorem and thus proving that, in "most cases", Hecke orbits contain no units at all.

We can investigate a mixed setting problem (multiplicative and modular), which is quite far from the original spirit of "avoiding problems", but which might be of interest.

Question. *Let $\sim$ be the smallest equivalence relation on $\overline{\mathbb{Q}}$ such that, for any $\alpha, \beta \in \overline{\mathbb{Q}}$,*

1. *if*
$$\mathbb{Q}(\alpha) = \mathbb{Q}(\beta)$$
then $\alpha \sim \beta$;

2. *if there are integers m, n, not both zero, such that*
$$\alpha^m = \beta^n$$
then $\alpha \sim \beta$;

3. *if there is a positive integer m such that*
$$\Phi_m(\alpha, \beta) = 0$$
then $\alpha \sim \beta$.

Can we find $\alpha, \beta \in \overline{\mathbb{Q}}$ such that $\alpha \not\sim \beta$?

If we forget about 1. and 2., the statement amounts to saying that the Hecke orbits are infinitely many. This is of course easy to prove, either exploiting bad reduction or using isogeny estimates.
One possible approach for 1., 2. and 3. simultaneously can be suggested by forgetting condition 3.; indeed, if α is rational and β is "related" to α via 1. and 2. only, we observe that β must be contained in a number field whose Galois group is solvable and hence no primitive element of a finite Galois extension of $\mathbb{Q}$ with Galois group, say, S_5, can be "related" to a rational.
If we have
$$\Phi_n(\alpha, \beta) = 0$$

then there is a tower of number fields $\mathbb{Q}(\alpha) \subseteq \mathbb{Q}(\alpha, \beta) \subseteq K$ with $\mathbb{Q}(\alpha) \subseteq K$ a Galois extension with Galois group isomorphic to a subgroup of $GL_2(\mathbb{Z}/n\mathbb{Z})$.
We might hope that, for n big enough, S_n cannot be "solved" using only quotients of subgroups of $GL_2(\mathbb{Z}/n\mathbb{Z})$ for varying n.

Bibliography

[1] Yves André, *Finitude des couples d'invariants modulaires singuliers sur une courbe algébrique plane non modulaire*, Journal für die reine und angewandte Mathematik **Issue 505** (1998).

[2] Yves André and Pietro Corvaja and Umberto Zannier and Ziyang Gao, *The Betti map associated to a section of an abelian scheme*, Inventiones mathematicae **222** (2020), no. 1, 161–202.

[3] Asvin G, *Unlikely and just likely intersections for high dimensional families of elliptic curves*, arXiv preprint arXiv:2203.16420 (2022).

[4] James Ax, *On Schanuel's conjectures*, Annals of mathematics **93** (1971), no. 2, 252–268.

[5] Gregorio Baldi and Bruno Klingler and Emmanuel Ullmo, *On the distribution of the Hodge locus*, Inventiones mathematicae **235** (2024), no. 2, 441–487.

[6] Francesco Ballini, *Division of primitive Points in an abelian Variety*, arXiv preprint arXiv:2207.00049 (2022).

[7] Alan Baker, *Transcendental number theory*, Cambridge University Press, 2022.

[8] Martin Bays and Philipp Habegger, *A note on divisible points of curves*, Transactions of the American Mathematical Society **367** (2015), no. 2, 1313–1328.

[9] Gal Binyamini and Christopher Daw, *Effective computations for weakly optimal subvarieties*, Journal of the European Mathematical Society (2024).

[10] Enrico Bombieri and David Masser and Umberto Zannier, *Intersecting a curve with algebraic subgroups of multiplicative groups*, International Mathematics Research Notices **1999** (1999), no. 20, 1119–1140.

[11] Laura Capuano and Peter Jossen and Christina Karolus and Francesco Veneziano, *Hyperelliptic continued fractions and generalized Jacobians: minicourse given by Umberto Zannier*, Arithmetic and Geometry: Ten Years in Alpbach (AMS) **202** (2019), 56.

[12] Chen Chung Chang and H Jerome Keisler, *Model theory*, Vol. 73, Elsevier, 1990.

[13] Laurent Clozel and Emmanuel Ullmo, *Equidistribution des points de Hecke*, Contributions to automorphic forms, geometry, and number theory (2004), 193–254.

[14] Pietro Corvaja and Julian Demeio and David Masser and Umberto Zannier, *On the torsion values for sections of an elliptic scheme*, Journal für die reine und angewandte Mathematik (Crelles Journal) **2022**, no. 782, 1–41.

[15] Sinnou David, *Minorations de formes linéaires de logarithmes elliptiques*, Société mathématique de France **62** (1995).

[16] _____, *Minorations de hauteurs sur les variétés abéliennes*, Bulletin de la Société Mathématique de France **121** (1993), no. 4, 509–544.

[17] Christopher Daw and Martin Orr, *Zilber-Pink in a product of modular curves assuming multiplicative degeneration*, arXiv preprint arXiv:2208.06338 (2022).

[18] Pierre Deligne, *Théorie de Hodge: II*, Publications Mathématiques de l'IHES **40** (1971), 5–57.

[19] Lou van den Dries and Chris Miller, *On the real exponential field with restricted analytic functions*, Israel Journal of Mathematics **85** (1994), no. 1-3, 19–56.

[20] Bas Edixhoven, *Special points on the product of two modular curves*, Compositio Mathematica **114** (1998), 307–320.

[21] Sebastian Eterović and Thomas Scanlon, *Likely intersections*, arXiv preprint arXiv:2211.10592 (2022).

[22] Gerd Faltings, *Endlichkeits-sätze für abelsche Varietäten über Zahlkörpern*, Invent. math **75** (1984), 38t.

[23] Éric Gaudron and Gaël Rémond, *Théorème des périodes et degrés minimaux d'isogénies*, Commentarii Mathematici Helvetici **89** (2014), no. 2, 343–403.

[24] Israel M Gelfand and Mikhail M Kapranov and Andrei V Zelevinsky, *Chow Varieties*, Springer, 1994.

[25] Samuel Grushevsky and Gabriele Mondello and Riccardo Salvati Manni and Jacob Tsimerman, *Compact Subvarieties of the Moduli Space of Complex Abelian Varieties*, arXiv preprint arXiv:2404.06009 (2024).

[26] Philipp Habegger, *Intersecting subvarieties of abelian varieties with algebraic subgroups of complementary dimension*, Inventiones mathematicae **176** (2009), no. 2, 405–447.

[27] ———, *On the bounded height conjecture*, International mathematics research notices **2009**, no. 5, 860–886.

[28] ———, *Singular moduli that are algebraic units*, Algebra & Number Theory **9** (2015), no. 7, 1515–1524.

[29] ———, *Weakly bounded height on modular curves*, Acta Math. Vietnam **35** (2010), no. 1, 43–69.

[30] Philipp Habegger and Jonathan Pila, *Some unlikely intersections beyond André-Oort*, Compositio Mathematica **148** (2012), no. 1, 1–27.

[31] Robin Hartshorne, *Algebraic geometry*, Vol. 52, Springer Science & Business Media, 2013.

[32] Marc Hindry and Joseph H Silverman, *Diophantine geometry: an introduction*, Vol. 201, Springer Science & Business Media, 2013.

[33] Monsur Akangbe Kenku, *On the number of Q-isomorphism classes of elliptic curves in each Q-isogeny class*, Journal of Number Theory **15** (1982), no. 2, 199–202.

[34] Bruno Klingler and Emmanuel Ullmo and Andrei Yafaev, *The hyperbolic Ax-Lindemann-Weierstrass conjecture*, Publications mathématiques de l'IHES **123** (2016), 333–360.

[35] Davide Lombardo, *Explicit open image theorems for abelian varieties with trivial endomorphism ring*, arXiv preprint arXiv:1508.01293 (2015).

[36] Julien Marché and Guillaume Maurin, *Singular intersections of subgroups and character varieties*, Mathematische Annalen **386** (2023), no. 1-2, 713–734.

[37] Guillaume Maurin, *Courbes algébriques et équations multiplicatives*, Mathematische Annalen **341** (2008), no. 4, 789–824.

[38] David Masser, *Linear relations on algebraic groups*, New Advances in Transcendence Theory (1988), 248–262.

[39] ———, *Small values of the quadratic part of the Néron-Tate height on an abelian variety*, Compositio Mathematica **53** (1984), no. 2, 153–170.

[40] David Masser and Jeffrey Vaaler, *Counting algebraic numbers with large height II*, Transactions of the American Mathematical Society **359** (2007), no. 1, 427–445.

[41] David Masser and Gisbert Wüstholz, *Estimating isogenies on elliptic curves*, Inventiones mathematicae **100** (1990), no. 1, 1-24.

[42] David Masser and Umberto Zannier, *Abelian varieties isogenous to no Jacobian*, Annals of Mathematics **191** (2020), no. 2, 635–674.

[43] ———, *Torsion anomalous points and families of elliptic curves*, American journal of mathematics **132** (2010), no. 6, 1677–1691.

[44] ———, *Torsion points on families of simple abelian surfaces and Pell's equation over polynomial rings (with an appendix by E. V. Flynn)*, Journal of the European Mathematical Society **17** (2015), no. 9, 2379–2416.

[45] Franz Mertens, *Ein Beitrag zur analytischen Zahlentheorie.* (1874).

[46] André Néron, *Quasi-fonctions et hauteurs sur les variétés abéliennes*, Annals of Mathematics (1965), 249–331.

[47] Martin Orr, *Height bounds and the Siegel property*, Algebra & Number Theory **12** (2018), no. 2, 455–478.

[48] Fabien Pazuki, *Modular invariants and isogenies*, International Journal of Number Theory **15** (2019), no. 03, 569–584.

[49] Jonathan Pila, *O-minimality and Diophantine geometry*, Proceedings ICM, 2014.

[50] ———, *On a modular Fermat equation*, Commentarii Mathematici Helvetici **92** (2017), no. 1, 85–103.

[51] ———, *On the algebraic points of a definable set*, Selecta Mathematica **15** (2009), no. 1, 151–170.

[52] ———, *Point-Counting and the Zilber-Pink Conjecture*, Cambridge University Press **228** (2022).

[53] Jonathan Pila and Ananth N Shankar and Jacob Tsimerman and Hélène Esnault and Michael Groechenig, *Canonical Heights on Shimura Varieties and the André-Oort Conjecture*, arXiv preprint arXiv:2109.08788 (2021).

[54] Jonathan Pila and Jacob Tsimerman, *Ax-Schanuel for the j-function*, Duke Mathematical Journal **165** (2016), no. 13, 2587–2605.

[55] Jonathan Pila and Alex James Wilkie, *The rational points of a definable set*, Duke Mathematical Journal **133(3)** (2006).

[56] Jonathan Pila and Umberto Zannier, *Rational points in periodic analytic sets and the Manin-Mumford conjecture*, Rendiconti Lincei **19** (2008), no. 2, 149–162.

[57] Alfred van der Poorten, *Power series representing algebraic functions*, Séminaire de théorie des nombres, Paris **91** (1990), 241–262.

[58] Jesus M Ruiz, *The basic theory of power series*, Springer, 1993.

[59] Stephen H Schanuel, *On heights in number fields*, Bulletin de la S. M. F. **107** (1979), 433-449.

[60] Stefan Schmidt, *Integrality Properties in the Moduli Space of Elliptic Curves: Isogeny Case*, The Quarterly Journal of Mathematics **73** (2022), no. 3, 1103–1136.

[61] Ananth N Shankar and Jacob Tsimerman, *Abelian varieties not isogenous to Jacobians over global fields*, arXiv preprint arXiv:2105.02998 (2021).

[62] Sergei Starchenko, *Notes on O-minimality*, 2009.

[63] Jacob Tsimerman, *Abelian varieties are not quotients of low-dimension Jacobians*, arXiv preprint arXiv:2302.05860 (2023).

[64] Emmanuel Ullmo, *Manin-Mumford, André-Oort, the equidistribution point of view*, Springer, 2007.

[65] Umberto Zannier, *Basics on the theory of heights and its applications to certain diophantine problems*, Expositiones Mathematicae **36** (2018), no. 1, 1–31.

[66] ———, *Hyperelliptic continued fractions and generalized Jacobians*, American Journal of Mathematics **141** (2019), no. 1, 1–40.

[67] ———, *Some Problems of Unlikely Intersections in Arithmetic and Geometry (AM-181)*, Princeton University Press, 2012.

[68] David Zywina, *Families of abelian varieties and large Galois images*, International Mathematics Research Notices **Issue 20** (2023), 17494–17551.