Introduction

An Overview of Some Problems of Unlikely Intersections

The present (rather long) introduction is intended to illustrate some basic problems of the topic, and to give an overview of the results and methods treated in the subsequent chapters. This follows the pattern I adopted in the lectures.

Let me first say a few words on the general title. This has to do with the simple expectation that when we intersect two varieties X, Y (whose type is immaterial now) of dimensions $r, s \geq 0$ in a space of dimension n, in absence of special reasons we expect the intersection to have dimension $\leq r+s-n$, and in particular to be empty if $r+s<n$. (This expectation may of course be justified on several grounds.)

More specifically, let X be fixed and let Y run through a denumerable set $\mathcal{Y}$ of algebraic varieties, chosen in advance independently of X, with a certain structure relevant for us, and such that $\dim X + \dim Y < n = \dim(\text{ambient})$; then we expect that only for a small subset of $Y \in \mathcal{Y}$, we shall have $X \cap Y \neq \emptyset$, unless there is a special structure relating X with $\mathcal{Y}$ which forces the contrary to happen. We shall usually express this by saying that X is a *special* variety.

When X is nonspecial, the said (expected) smallness may be measured in terms of the union of the intersections $\bigcup_{Y \in \mathcal{Y}}(X \cap Y)$: how is this set distributed in X? Is this set finite?

Similarly, we may study analogous situations when $\dim Y = s$ is any fixed number, whereas $\dim(X \cap Y) > \dim(X) + s - n$ for several $Y \in \mathcal{Y}$.

We note that often these problems can also be seen as expressing some kind of *local-global* principle: a point of intersection of X with some $Y \in \mathcal{Y}$ encodes a local property of a suitable set of coordinate functions on X at that point; we expect this property to occur only at a few points, unless it is the specialization of a global property of these functions on X. Such a global property should correspond to X being "special."

Now, it turns out that some known problems involving arithmetic and geometry can be put into this (rough) context, where the varieties in $\mathcal{Y}$ are usually described by equations of growing degrees, and depending on discrete parameters.[1] For instance, $\mathcal{Y}$ could consist of denumerably

[1] So, especially because the degree (over $\mathbb{Q}$) grows, the varieties in $\mathcal{Y}$ considered here do not vary in *families* in the common algebraic or continuous meaning. In particular, the said intersections (whether unlikely or not) usually cannot be described by (rational or integral) points on algebraic varieties. This fact introduces an arithmetical aspect in the problems seemingly of different type compared to the usual diophantine questions. So, for instance, although a line meeting a given curve in $\mathbb{A}^3$ may be considered to produce an "unlikely intersection," the diophantine issue of describing the set of such lines which are defined over $\mathbb{Q}$ does not fall in the realm considered here.

many prescribed points, which is indeed the case in many of the basic issues in this topic; such points shall be called the *special points*. In each case, the special varieties shall constitute the natural (for the structure in question) higher-dimensional analogue of the special points. Then we expect that a nonspecial variety X of positive codimension shall contain only a few special points, for instance a set which is not Zariski-dense.

Let me give a simple example at the basis of the problems to be discussed.

Lang's problem on roots of unity

Such an issue was raised by S. Lang in the 1960s. He posed the following attractive problem, a kind of simple prototype of other questions we shall touch: *suppose that $X : f(x, y) = 0$ is a complex plane irreducible curve containing infinitely many points (ζ, θ) whose coordinates are roots of unity; what can be said of the polynomial f?*

Actually, equations in roots of unity go back to long ago: P. Gordan already in 1877 studied certain equations of this type, linear and with rational coefficients, related to the classification of finite groups of homographies. In part inspired by this and by subsequent papers, e.g., of H.B. Mann, the subject was also investigated in a systematic way by J.H. Conway and A.J. Jones [CJ76]. In their terminology, we may view such problems as *trigonometric diophantine equations*; in fact, if we write the coordinates in exponential shape, we have a trigonometric equation $f(\exp(2\pi i\alpha), \exp(2\pi i\beta)) = 0$, to be solved in rational "angles" $\alpha, \beta \in \mathbb{Q}/\mathbb{Z}$.[2]

Observe also that the points with roots of unity coordinates are precisely the *torsion points* in the algebraic group $\mathbb{G}_{\mathrm{m}}^2$. (We recall that as a variety $\mathbb{G}_{\mathrm{m}}$ is simply $\mathbb{A}^1 \setminus \{0\}$, the affine line with the origin removed; we endow it with the multiplicative group law to make it into an algebraic group.) The torsion points constitute the set $\mathcal{Y}$ of special points in this problem.

Lang actually expected only finitely many torsion points to lie in X, unless a special (multiplicative) structure occurred, which he formulated as X being a *translate of an algebraic subgroup by a torsion point*, which we call a *torsion coset*. This amounts to the equation $f(x, y) = 0$ being (up to a monomial factor) of the shape $x^a y^b = \rho$, for integers a, b not both zero (their sign is immaterial here) and ρ a root of unity. This structure is actually clearly unavoidable because it yields infinitely many torsion points in X. We call it the *special structure* for the problem in question, and we call the torsion cosets the *special irreducible (sub)varieties* for this issue; they are also named *torsion varieties*.

The result foreseen by Lang can be rephrased by stating that *an irreducible curve contains infinitely many special (=torsion) points if and only if it is a special (=torsion) curve.*

As mentioned in [Lan83], this expectation of Lang was soon proved by Ihara, Serre, Tate (see next chapter and also [Lan65] for an account of these proofs); it was accompanied by other questions (also of others), such as what happens in higher dimensions and for other algebraic groups, and provided further motivation for them.

Let us give a description of these evolutions and of some other related issues, which will serve also as a sort of summary for the topics of these notes. They involve several different methods, of which I shall describe only a small part in some detail.

Summary

Chapter 1: *Unlikely Intersections in Multiplicative Groups and the Zilber Conjecture.*

[2]The quoted paper of Conway-Jones constructs a theory which reduces trigonometric diophantine equations to usual diophantine equations. It also mentions a number of classical applications, including the one noted by Gordan (1877), of the problem of solving equations in roots of unity.

We shall discuss unlikely intersections in (commutative) multiplicative algebraic groups (over a field of characteristic zero). For our purposes we may perform a finite extension of the ground field, and then such groups are of the shape $\mathbb{G}_{\mathrm{m}}^n$; they are also called (algebraic) tori.

The above-mentioned Lang's problem is the simplest nontrivial issue in this context. The natural generalization to higher dimensions also formed the object of a conjecture of Lang, proved by M. Laurent [Lau84] and independently by Sarnak-Adams [SA94]. The final result, of which we shall sketch a proof (by a number of methods) in Chapter 1, may be stated in the following form:

Theorem. *Let Σ be a set of torsion points in $\mathbb{G}_{\mathrm{m}}^n(\overline{\mathbb{Q}})$. The Zariski closure of Σ is a finite union of torsion cosets.*

The torsion cosets are by definition the translates of algebraic subgroups by a torsion point; it turns out that they may be always defined by finitely many equations of the shape $x_1^{a_1} \cdots x_n^{a_n} = \theta$, for integers a_i and root of unity θ. They are the special varieties in this context; in the case when the dimension is 0 we find the special points, i.e., the torsion points. Then, a rephrasing is that

The Zariski closure of a(ny) set of special points is a finite union of special varieties.

Note that if we start with any (irreducible) algebraic variety $X \subset \mathbb{G}_{\mathrm{m}}^n$ and take Σ as the set of all torsion points in X, we find that Σ is confined to a finite union of torsion cosets in X, and hence is not Zariski-dense in X unless X is itself a torsion coset.

This result in practice describes the set of solutions of a system of algebraic equations in roots of unity, and confirms the natural intuition that all the solutions are originated by a multiplicative structure of finitely many subvarieties of X. As noted above, such solutions in roots of unity had been treated from a somewhat different viewpoint also by Mann [Man65] and Conway-Jones [CJ76].

These conjectures and results inspired analogous and deeper problems; for instance, we may replace $\mathbb{G}_{\mathrm{m}}^n$ by an abelian variety and ask about the corresponding result. In the case of a complex curve X of genus at least 2 embedded in its Jacobian, such a problem was raised independently by Yu. Manin and separately by D. Mumford, already in the 1960s; it predicted finiteness for the set of torsion points on X. Actually, Mumford's question apparently motivated in part Lang's above problem, as mentioned in [Lan65]; then Lang was led to unifying statements. The Manin-Mumford conjecture was proved by M. Raynaud, who soon was able to analyze completely the general case of an arbitrary subvariety X of an abelian variety A (see, e.g., [Ray83]).

The final result by Raynaud may be phrased in the above shape:

Raynaud's Theorem. *Let A be an abelian variety defined over a field of characteristic* 0 *and let Σ be a set of torsion points in A. Then the Zariski closure of Σ is a finite union of translates of abelian subvarieties of A by a torsion point.*

The fundamental case occurs when A is defined over $\overline{\mathbb{Q}}$, which indeed implies (by specialization) the general case of characteristic zero, whereas the conclusion is generally false, e.g., in the case of $\overline{\mathbb{F}}_p$, since any point is then torsion. The special varieties of this abelian context are the torsion-translates of abelian subvarieties, and the special points are again the torsion points.

There are now several known proofs of the Manin-Mumford conjecture and its extensions, but none of them is really easy, and the matter is distinctly deeper than the toric case. A new proof appears in the paper [PZ08] with a method which applies also to other issues on unlikely intersections where other arguments do not apply directly. This method and its implications constitute one of the main topics we shall discuss, however, in Chapters 3 and 4.

This context evolved in several deep directions, first with the study of points in $X \cap \Gamma$ for a subgroup Γ of finite rank (Lang, P. Liardet, M. Laurent, G. Faltings, P. Vojta...), or later with the study of algebraic points of small height (after F. Bogomolov); actually, certain issues on "small points" implicitly motivated some of the studies we shall discuss later. For the sake of completeness, in Chapter 1 we shall briefly mention a few results on this kind of problem, without pausing, however, on any detail.

Still in another direction, we may continue the above problems to higher multiplicative rank, by intersecting a given variety $X \subset \mathbb{G}_m^n$ not merely with the set of torsion points, but with the family of algebraic subgroups of $\mathbb{G}_m^n$ up to any given dimension; when this dimension is 0, we find back the torsion points.

The structure of algebraic subgroups of $\mathbb{G}_m^n$ (recalled below) shows that a point $(u_1, \ldots, u_n) \in X$ lies in such a subgroup of dimension r if and only if the coordinates $u_1, \ldots, u_n$ satisfy at least $n - r$ independent relations $u_1^{m_1} \cdots u_n^{m_n} = 1$ of multiplicative dependence, i.e., the rank of the multiplicative group $u_1^{\mathbb{Z}} \cdots u_n^{\mathbb{Z}}$ is at most r.

We find maximal dependence when the point is torsion, i.e., we may take $r = 0$. But of course the intersection shall be unlikely as soon as $r + \dim X < n$. So, under this condition, we already expect a sparse set of intersections, unless X is "special" in some appropriate algebraic sense. For instance, we do not expect X to be special if the coordinates $x_1, \ldots, x_n$ are multiplicatively independent as functions on X; in this case we should expect this independence to be preserved by evaluation at most points of X. We clearly see here a kind of local-global principle alluded to above.

A prototype of this kind of problem was studied already by Schinzel in the 1980s (see, e.g., Ch. 4 of [Sch00]) in the course of his theory of irreducibility of lacunary polynomials, so with independent motivations; he obtained fairly complete results only when X is a curve in a space up to dimension $n \leq 3$. For the case of arbitrary dimension n, confining again to *curves* X, this study was the object of a joint paper with E. Bombieri and D. Masser [BMZ99], and then was studied further by P. Habegger, G. Rémond, G. Maurin, and also by M. Carrizosa, E. Viada, and others in the case of abelian varieties (here the results are less complete).

We shall present in some detail the two main results of the paper [BMZ99] in Chapter 1. The common assumption is that the curve X is not contained in a translate of a proper algebraic subgroup of $\mathbb{G}_m^n$; this means that the coordinate functions $x_1, \ldots, x_n$ are multiplicatively independent modulo constants. Under this assumption we have (working over $\overline{\mathbb{Q}}$):

Theorem 1. *If X is an irreducible curve over $\overline{\mathbb{Q}}$, not contained in any translate of a proper algebraic subgroup of $\mathbb{G}_m^n$, then the Weil height in the set $\bigcup_{\dim G \leq n-1} (X \cap G)$ is bounded above.*[3]

Here and below, G is understood to run through algebraic subgroups of $\mathbb{G}_m^n$. Note that the sum $\dim X + \dim G$ here may be equal to n, so these intersections $X \cap G$ may be indeed considered "likely intersections"; actually, it is easily proved that they constitute an infinite set. However, the result shows that they are already sparse: in fact, for instance the well-known (easy) Northcott's theorem immediately implies that there are only finitely many such points of bounded degree. On the other hand, note also that a priori it is not even clear that these intersections do not exhaust all the algebraic points on our curve X.

In particular, Theorem 1 applies to the curve $x_1 + x_2 = 1$, predicting that such x_1, x_2 which are multiplicatively dependent have bounded height; this example, first proposed by Masser (and analogue of an example by S. Zhang and D. Zagier for lower bounds for heights), played a motivating role in the early work.

To go on, let us now look at algebraic subgroups G with $\dim G \leq n - 2$; now we impose at least two multiplicative relations, and we have, so to say, double sparseness, and truly unlikely intersections. This is confirmed by the following result:

Theorem 2. *If X is an irreducible curve over $\overline{\mathbb{Q}}$, not contained in any translate of a proper algebraic subgroup of $\mathbb{G}_m^n$, then $\bigcup_{\dim G \leq n-2} (X \cap G)$ is finite.*

[3] We shall briefly recall this notion of height in Chapter 1.

For instance, the set A (resp. B) of algebraic numbers x (resp. y) such that $x, 1-x$ (resp. $y, 1+y$) are multiplicatively dependent has bounded height, by Theorem 1, whereas $A \cap B$ is finite, by Theorem 2 (applied to the line in $\mathbb{G}_{\mathrm{m}}^3$ parametrized by $(t, 1-t, 1+t)$).

We shall give sketches of proofs of these theorems, which roughly speaking depend on certain comparisons between degrees and heights.

The assumption that X is not contained in a *translate* of a proper algebraic subgroup, rather than just in a proper algebraic subgroup (or, equivalently, in a *torsion-translate* of a proper algebraic subgroup) makes a subtle difference; it is necessary for the first result to hold, but this necessity was not clear for the second one. This turned out to be in fact an important issue, because for instance it brought into the picture the deep Lang's conjectures (alluded to above) on the intersections (of a curve) with finitely generated groups. Only recently has it been proved by G. Maurin [Mau08] that the weaker assumption suffices.

Before Maurin's proof, the attempt to clarify this issue (as, for instance, in [BMZ06]) led to other natural and independent questions, like an extension of Theorem 1 to higher dimensional varieties; it was soon realized that for this aim new assumptions were necessary, and in turn this led to the consideration of unlikely intersections of higher dimensions. Such a study was performed in [BMZ07], where among other things a kind of function field analogue of Theorem 2 was obtained, for unlikely intersections of positive dimension (with algebraic cosets); also, the issue of the height was explicitly stated therein with a "bounded height conjecture." This was eventually proved by P. Habegger in [Hab09c] with his new ideas. With the aid of this result, a new proof of Maurin's theorem was also achieved in [BHMZ10].

In the meantime, after the paper [BMZ99] was published, it turned out that quite similar problems had been considered independently and from another viewpoint also by B. Zilber, who, with completely different motivations arising from model theory, had formulated in [Zil02] general conjectures for varieties of arbitrary dimensions (also in the abelian context), of which the said theorem of [Mau08] is a special case. Another independent formulation of such conjectures (even in greater generality) was given by R. Pink (unpublished). These conjectural statements contain several of the said theorems[4] and in practice predict a certain natural finite description for all the unlikely intersections in question.

We shall state Zilber's conjecture (in the toric case) and then present in short some extensions of the above theorems, some other results (e.g., on the said unlikely intersections of positive dimension) and some applications, for instance, to the irreducibility theory of lacunary polynomials. We shall also see how Zilber's conjecture implies uniformity in quantitative versions of the said results.

Further, in the notes to the chapter we shall offer some detail about an independent method of Masser to study the sparseness of the intersections considered in Theorem 1 (actually with a milder assumption), and we shall see other more specific questions.

Chapter 2: *An Arithmetical Analogue*

The unlikely intersections of the above Theorem 2 for X a curve in $\mathbb{G}_{\mathrm{m}}^n$ correspond to (complex) solutions to pairs of equations $\mathbf{x}^{\mathbf{a}} = \mathbf{x}^{\mathbf{b}} = 1$ on X, where we have abbreviated, e.g., $\mathbf{x}^{\mathbf{a}} := x_1^{a_1} \cdots x_n^{a_n}$. The x_i are the natural coordinate functions on $\mathbb{G}_{\mathrm{m}}^n$, whereas $\mathbf{a}, \mathbf{b}$ vary over all pairs of linearly independent integral vectors. In other words, we are considering common zeros of two rational functions $u-1$ and $v-1$ on X, where both $u := \mathbf{x}^{\mathbf{a}}$, $v := \mathbf{x}^{\mathbf{b}}$ are taken from a finitely generated multiplicative group Γ of rational functions, namely, $\Gamma = x_1 \cdots x_n$, the group generated by the coordinates $x_1, \ldots, x_n$.

[4] However, they do not consider heights.

In this view, we obtain an analogue issue for number fields k on considering a finitely generated group Γ in k^* (e.g., the group of S-units, for a prescribed finite set S of places) and on looking at primes dividing both $u-1$ and $v-1$, for u, v running through Γ. These primes now constitute the unlikely intersections; note that now there are always infinitely many ones (if Γ is infinite).

For *given* u, v, a measure of the magnitude of the set of these primes is the $\gcd(1-u, 1-v)$. It turns out that if u, v are multiplicatively independent this can be estimated nontrivially; for instance, we have:

<u>Theorem</u>: *Let $\epsilon > 0$ and let $\Gamma \subset \mathbb{Q}^*$ be a finitely generated subgroup. Then there is a number $c = c(\epsilon, \Gamma)$ such that if $u, v \in \mathbb{Z} \cap \Gamma$ are multiplicatively independent, we have* $\gcd(1-u, 1-v) \leq c \max(|u|, |v|)^\epsilon$.

We shall give a proof of this statement, relying on the subspace theorem of Schmidt, which we shall recall in a simplified version that is sufficient here.

A substantial difference with the above context is that here we estimate the individual intersections, whereas previously we estimated their union over all pairs $u, v \in \Gamma$. However, in this context a uniform bound for the union does not hold.

In the special cases $u = a^n, v = b^n$ (with a, b fixed integers, $n \in \mathbb{N}$), the displayed results were obtained in joint work with Y. Bugeaud and P. Corvaja [BCZ03], whereas the general case was achieved in [CZ03] and [CZ05], also for number fields (where the gcd may be suitably defined, also involving archimedean places). All of these proofs use the above-mentioned Schmidt subspace theorem (which is a higher-dimensional version of Roth's theorem in diophantine approximation; see, e.g., [BG06] for a proof and also for its application to the present theorem).

These results admit, for instance, an application to the proof of a conjecture by Gyory-Sarkozy-Stewart (that the greatest prime factor of $(ab+1)(ac+1) \to \infty$ as $a \to \infty$, where $a > b > c > 0$). They also have applications to various other problems, including the structure of the groups $E(\mathbb{F}_{q^n})$, $n \to \infty$, for a given ordinary elliptic curve $E/\mathbb{F}_q$ (Luca-Shparlinski [LS05]) and to a Torelli Theorem over finite fields (Bogomolov-Korotiaev-Tschinkel, [BT08] and [BKT10]).

J. Silverman [Sil05] formulated the result in terms of certain heights and recovered it as a consequence of a conjecture of Vojta for the blow-up of $\mathbb{G}_m^2$ at the origin; we shall recall in brief these interpretations.

There are also analogous estimates over function fields (obtained in [CZ08b]); they also provide simplification for some proofs related to results in Chapter 1, which yields further evidence that such an analogy is not artificial. These estimates also have implications to the proof of certain special cases of Vojta's conjecture on integral points over function fields, specifically for $\mathbb{P}_2 \setminus$ three divisors, and for counting rational points on curves over finite fields. We shall present (also in the notes) some of these results and provide detail for some of the proofs.

Chapter 3: *Unlikely Intersections in Elliptic Surfaces and Problems of Masser*

D. Masser formulated the following attractive problem. Consider the Legendre elliptic curve $E_\lambda : Y^2 = X(X-1)(X-\lambda)$ and the points $P_\lambda = (2, \sqrt{2(2-\lambda)})$, $Q_\lambda = (3, \sqrt{6(3-\lambda)})$ on it. Here λ denotes an indeterminate, but we can also think of specializing it. Consider then the set of complex values $\lambda_0 = 0, 1$ of λ such that both P_{λ_0} and Q_{λ_0} are torsion points on E_{λ_0}. It may be easily proved that none of the points is identically torsion on E_λ and that actually the points are linearly independent over $\mathbb{Z}$, so it makes sense to ask:

Masser's problem: *Is this set finite?*

Clearly this is also a problem of unlikely intersections: for varying $\lambda \in \mathbb{C}$, the E_λ describe a(n elliptic) surface, with a rational map λ to $\mathbb{P}_1$. The squares E_λ^2, again for varying $\lambda = 0, 1$, describe a threefold, i.e., the fibered product (with respect to the map λ) $E_\lambda \times_{\mathbb{P}_1\setminus\{0,1,\infty\}} E_\lambda$. (This is an elliptic group-scheme over $\mathbb{P}_1 \setminus \{0,1,\infty\}$.) The $P_\lambda \times E_\lambda$ and $E_\lambda \times Q_\lambda$ both describe a surface, whereas the points $P_\lambda \times Q_\lambda$ describe a curve in this threefold; also, each condition $mP_\lambda = O_\lambda$ (i.e., the origin of E_λ) or $nQ_\lambda = O_\lambda$ also corresponds to a surface (if $mn = 0$), whereas a pair of such conditions yields a curve, since the points are linearly independent on E_λ (as we shall see).

Putting together these dimensional data, we should then expect

(i) that each point gives rise to infinitely many "torsion values" λ_0 of λ (which may be proved, however, a bit less trivially than might be expected), but

(ii) that it is unlikely that a same value λ_0 works for both points, which would correspond to an intersection of two curves in the threefold.

Hence, a finiteness expectation in Masser's context is indeed sensible.

Of course, this is just a simple example of analogous questions that could be stated. For instance, we could pick any two points in $E_\lambda(\overline{\mathbb{Q}(\lambda)})$ and ask the same question. In the general case we would expect finiteness only when the points are linearly independent in E_λ; any (identical) linear dependence would correspond to a *special variety* on this issue of unlikely intersections.

In fact, it turned out that S. Zhang had raised certain similar issues in 1998, and R. Pink in 2005 raised independently such a type of conjecture in the general context of semiabelian group schemes, generalizing the present problems. In these notes we shall stick mainly to special cases like the one above, but we shall discuss also some recent progress toward more general cases.

Note that if we take a fixed "constant" elliptic curve E, say over $\overline{\mathbb{Q}}$, in place of E_λ, and if we take two points $P, Q \in E(\overline{\mathbb{Q}(\lambda)})$, the obvious analogue of Masser's question reduces to the Manin-Mumford problem for E^2, namely, we are just asking for the torsion points in the curve described by $P_\lambda \times Q_\lambda$ in E^2 (the fact that now E is constant allows us to work with the surface E^2 rather than the threefold $E^2 \times \mathbb{P}_1$). The *special varieties* here occur when the points are linearly dependent: $aP_\lambda = bQ_\lambda$ for integers a, b not both zero.[5]

However, the known arguments for Raynaud's theorem seem not to carry over directly to a *variable* elliptic curve as in Masser's problem.

In Chapter 3, we shall actually start with the Manin-Mumford conjecture (in extended form, i.e., Raynaud's theorem), sketching the mentioned method of [PZ08] to recover it, and we shall then illustrate in some detail how the method also applies in this relative situation. The method proves the finiteness expectation, as in the papers [MZ08] and [MZ10b]. (In more recent work [MZ10d], it is shown that this method suffices also for any choice of two points on E_λ, with coordinates in $\overline{\mathbb{Q}(\lambda)}$ and not linearly dependent.)

The principle of the method is, very roughly, as follows. Let us consider for simplicity the case of the Manin-Mumford issue of torsion points on a curve X in a fixed abelian variety $A/\overline{\mathbb{Q}}$. We start by considering a transcendental uniformization $\pi : \mathbb{C}^g/\Lambda \to A$, where Λ is a full lattice in $\mathbb{C}^g$. By means of a basis for Λ, we may identify $\mathbb{C}^g$ with $\mathbb{R}^{2g}$, and under this identification the torsion points on A (of order N) become the rational points in $\mathbb{R}^{2g}$ (of denominator N).

Thus the torsion points on X of order N give rise to rational points on $Z := \pi^{-1}(X)$ of denominator N; in the above identification, Z becomes a real-analytic surface. Now the proof compares two kind of estimates for these rational points:

• Upper bounds: By work of Bombieri-Pila [BP89], generalized by Pila [Pil04] and further by Pila-Wilkie[PW06], one can often estimate nontrivially the number of rational points with denominator dividing N on a (compact part of a) transcendental variety Z;[6] the estimates take

[5] If E has complex multiplication, we should, however, take into account these relations with $a, b \in \mathrm{End}(E)$.

[6] The original paper [BP89] strongly influenced the subsequent work; it was concerned with curves, also algebraic, and the attention to transcendental ones came also from a specific question of Sarnak.

the shape $\ll_{\epsilon,Z} N^{\epsilon}$, for any $\epsilon > 0$, *provided*, however, we remove from Z the union of connected semialgebraic arcs [7] contained in it. (This proviso is necessary, as these possible arcs could contain many more rational points.) In the present context, purely geometrical considerations show that if X is not a translate of an elliptic curve, then Z does not contain any such semialgebraic arc. Thus the estimate holds indeed for all rational points on Z.

• Lower bounds: Going back to the algebraic context of A and X, we observe that a torsion point $x \in X$ carries all its conjugates $x^{\sigma} \in X$ over a number field k of definition for X; these conjugates are also torsion, of the same order as x. Further, by a deep estimate of Masser [Mas84] (coming from methods stemming from transcendence theory), their number (i.e., the degree $[k(x) : k]$) is $\gg_A N^{\delta}$, where $\delta > 0$ is a certain positive number depending only on (the dimension of) A.

Conclusion: Plainly, comparison of these estimates proves (on choosing $\epsilon < \delta$) that N is bounded unless X is an elliptic curve inside A, as required.

This method would work also for the simpler context of torsion points in subvarieties of $\mathbb{G}_{\mathrm{m}}^n$, as in the original question by Lang. (See, e.g., [Sca11b].)

In the case of Masser's above-mentioned problem, things look somewhat different, since we have a family of elliptic curves E_λ, and then the uniformizations $\mathbb{C}/\Lambda_\lambda \to E_\lambda$ make up a family as well. Also, when we take conjugates of a torsion point in E_{λ_0}, we fall in other curves $E_{\lambda_0^\sigma}$. However, in spite of these discrepancies with the "constant" case, it turns out that the issue is still in the range of the method: we may use a varying basis for the lattices Λ_λ (constructed by hypergeometric functions) and again define real coordinates which take rational values on torsion points. And then the above-sketched proof-pattern still works in this relative picture. We shall present a fairly detailed account of this.

Further problems: Masser formulated also analogous issues, e.g., for other abelian surfaces and further for a larger number of points (depending on more parameters). There are also interesting applications, such as an attractive special case considered by Masser, concerning polynomial solutions in $\mathbb{C}[t]$ of Pell's equations, e.g., of the shape $x^2 = (t^6 + t + \lambda)y^2 + 1$.

Independently of this, Pink also stated related and more general conjectures for group-schemes over arbitrary varieties. In Chapter 3 we shall also mention some of these possible extensions, and some more recent work.

The said method in principle applies in greater generality, but in particular it often needs as a crucial ingredient a certain height bound, which is due to Silverman for the case of a single parameter. In higher dimensions, an analogous bound was proved only recently by P. Habegger, which may lead to significant progress toward the general questions of Masser and Pink.

A dynamical analogue. It is worth mentioning an analogue of these questions in algebraic dynamics; this comes on realizing that the torsion points on an elliptic curve correspond to the preperiodic points for the so-called Lattès map, namely the rational map $\mathbb{P}_1 \to \mathbb{P}_1$ of degree 4 which expresses $x(2P)$ in terms of $x(P)$, for a point $P \in E$, where x is the first coordinate in a Weierstrass model. (For the Legendre curve this map is $(x^2 - \lambda)^2/4x(x-1)(x-\lambda)$.) A reformulation in this context of Masser's question becomes:

Are there infinitely many $\lambda \in \mathbb{C}$ such that $2, 3$ are both preperiodic for the Lattès map x_λ relative to E_λ?

At the AIM meeting in Palo Alto (Jan. 2008) I asked whether such a question may be dealt with for other rational functions depending on a parameter, for instance $x^2 + \lambda$. M. Baker and L. DeMarco recently succeeded, in fact, in proving [BD11] that, if $d \geq 2$, then $a, b \in \mathbb{C}$ are preperiodic

[7] By "semialgebraic arc" we mean the image of a C^∞ nonconstant map from an interval to a real algebraic curve.

with respect to $x^d + \lambda$ for infinitely many $\lambda \in \mathbb{C}$, if and only if $a^d = b^d$ (this is the *special variety* in this problem of unlikely intersections).

The method of Baker-DeMarco is completely different from the one above, and we shall only say a few words about it; it is likely that it applies to the Lattès maps as well (although this has not yet been carried out). In any case it provides a complement to the method used by Masser, Pila, and myself.

We shall conclude by mentioning and discussing as well some dynamical analogues of the Manin-Mumford conjecture.

Chapter 4: *About the André-Oort Conjecture*

Recently, J. Pila found that the method that we have very briefly described in connection with Chapter 3, on the problems of Masser and the conjectures of Manin-Mumford and Pink, can be applied to another bunch of well-known problems that go under the name "André-Oort conjecture."

This conjecture concerns Shimura varieties and is rather technical to state in the most general shape; fundamental examples of Shimura varieties are the modular curves[8] and, more generally, the moduli spaces parametrizing principally polarized abelian varieties of given dimension, possibly with additional (level) structure (see below for simple examples and Chapter 4 for more). At this point let us merely say that the pattern of the relevant statement is similar to the others we have found; in fact, there is a notion of Shimura subvariety of a Shimura variety, and if we interpret these subvarieties as being the "special" ones (in dimension 0 we find the special points), the general statement becomes:

André-Oort conjecture: *The Zariski closure of a set of special points in a Shimura variety is a special subvariety,*

or, equivalently:

If a subvariety of a Shimura variety has a Zariski-dense set of special points, then it is a special subvariety.

For the case when the said subvariety is a curve, this statement first appeared (in equivalent form) as Problem 1 on p. 215 of Y. André's book [And89]; then it was stated independently by F. Oort at the Cortona Conference in 1994 for the case of moduli spaces of principally polarized abelian varieties of given dimension [Oor97]. (We also note that in [And89], p. 216, André already pointed out a similarity with the Manin-Mumford conjecture.)

To illustrate the simplest instances of this conjecture, let me briefly recall a few basic facts in the theory of elliptic curves (see, e.g., [Sil92]). Every complex elliptic curve E may be defined (in $\mathbb{A}^2$) by a Weierstrass equation $E : y^2 = x^3 + ax + b$, where $a, b \in \mathbb{C}$ are such that $4a^3 + 27b^2 = 0$. It has an invariant $j = j(E) := 1728\frac{4a^3}{4a^3+27b^2}$, which is such that E, E' are isomorphic over $\mathbb{C}$ if and only if they have the same invariant.

Hence we may view the affine (complex) line $\mathbb{A}^1$ as parametrizing (isomorphism classes of) elliptic curves through the j-invariant; this is the simplest example of modular curves, which in turn are the simplest Shimura varieties of positive dimension. Other modular curves are obtained as finite covers of $\mathbb{A}^1$, parametrizing elliptic curves plus some discrete additional structure, for instance a choice of a torsion point of given order, or a finite cyclic subgroup of given order.

A "generic" complex elliptic curve has an endomorphism ring equal to $\mathbb{Z}$; however, it may happen that the endomorphism ring is larger, in which case it is an *order* in the ring of integers of an imaginary quadratic field (and any such order is possible). In such cases the invariant j

---

[8] These are algebraic curves, with an affine piece analytically isomorphic to $\Gamma\backslash\mathcal{H}$, where $\mathcal{H}$ is the upper-half-plane and Γ is a so-called congruence subgroup of $SL_2(\mathbb{Z})$.

turns out to be an algebraic number (actually algebraic integer), and equal to a value $j(\tau)$ of the celebrated "modular function" $j(z)$ at some imaginary quadratic number τ in the upper-half plane $\mathcal{H} := \{z \in \mathbb{C} : \Im z > 0\}$. These values are called "singular (moduli)" or "CM" (from "complex multiplication," which is the way in which the endomorphisms arise). They turn out to be the "special" points in the modular curves (and they are the smallest Shimura varieties, of dimension 0). We shall be more detailed on this in Chapter 4, with an explicit section devoted to a brief review of all of this; see also [Lan73], [Shi94], [Sil92] and [Sil02].

Now, the André-Oort conjecture for (the modular) curves is a trivial statement, since there are no varieties strictly intermediate between points and an irreducible curve. However, if we take the product of two modular curves, such as $\mathbb{A}^1 \times \mathbb{A}^1 = \mathbb{A}^2$, we may choose X as any (irreducible) plane curve in this product, as an intermediate variety. To see the shape of the conjecture in this context, let us briefly discuss the special points and the special curves in $\mathbb{A}^2$. The special points in $\mathbb{A}^2$ are the pairs of special points in $\mathbb{A}^1$. The special curves turn out to be of two types. A first (trivial) type is obtained by taking $\{x\} \times \mathbb{A}^1$ or $\mathbb{A}^1 \times \{x\}$, where $x \in \mathbb{A}^1$ is a special point. A second type is obtained by the modular curves denoted $Y_0(n)$; such a curve is defined by a certain symmetric (if $n > 1$) irreducible polynomial $\Phi_n(x_1, x_2) \in \mathbb{Q}[x_1, x_2]$, constructed explicitly, e.g., in [Lan73], Ch. 5. The curve $Y_0(n)$ parametrizes (isomorphism classes of) pairs (E_1, E_2) of elliptic curves such that there exists a cyclic isogeny $\phi : E_1 \to E_2$ of degree n. We note that all of these curves contain infinitely many special points; for the first type this is clear, whereas for the second type it suffices to recall that isogenous curves have isomorphic fields of endomorphisms.[9]

Coming back to our (irreducible) plane curve $X \subset \mathbb{A}^2$, note that we may view any point $(x_1, x_2) \in X$ as corresponding to a pair of (isomorphism classes of) elliptic curves E_1, E_2 with $j(E_i) = x_i$ for $i = 1, 2$. This point is special if E_1, E_2 are both CM-elliptic curves. Note that we may fix the coordinate x_1 as we please, but of course this shall determine x_2 up to boundedly many possibilities, so if x_1 is a special point, it shall be unlikely that some suitable x_2 shall also be special.

The André-Oort conjecture in this setting states that *there are infinitely many special points on the irreducible curve $X \subset \mathbb{A}^2$ precisely if X is of one of the two types we have just described above.*

This basic case of the conjecture is the precise analogue for this context of the question of Lang about torsion points on a plane curve, which we have recalled at the beginning, and which provided so much motivation for subsequent work. In this special case the conjecture was proved by André himself in [And98]. (The proof readily generalizes to arbitrary products of two modular curves.) The conclusion may be viewed as the complete description of the fixed algebraic relations holding between infinitely many pairs of CM-invariants. (An attractive example concerns the Legendre curves $E_\lambda : y^2 = x(x-1)(x-\lambda)$. The result implies the finiteness of the set of complex λ_0 such that both E_{λ_0} and $E_{-\lambda_0}$ have complex multiplication.)

We shall reproduce this proof by André in Chapter 4 (together with an effective variation: see the remarks below).

Another argument for André's theorem, by B. Edixhoven, appeared almost simultaneously [Edi98]; however, this was conditional to the Generalized Riemann Hypothesis for zeta-functions of imaginary quadratic fields. A merit of this argument is that it opened the way to generalizations; for instance, Edixhoven treated arbitrary products of modular curves [Edi05], and other extensions were obtained by Edixhoven and A. Yafaev [EY03]. Recently this has also been combined with equidistribution methods (by L. Clozel and E. Ullmo), and a *conditional* proof of the full conjecture

[9] We mean fields $\mathrm{End}(E) \otimes \mathbb{Q}$.

has been announced recently by B. Klingler, Ullmo, and Yafaev. (See also [Noo06] and [Pil11] for other references.) We shall briefly comment on Edixhoven's method in the notes to Chapter 4.

More recently Pila in [Pil09b] succeeded in applying *unconditionally* the method we have mentioned while summarizing Chapter 3 (at the basis of the papers [MZ08], [PZ08], and [MZ10b]) to these problems in the André-Oort context. In place of the uniformization of an abelian variety by a complex torus, he used the uniformization of modular curves provided by the modular function.

He first obtained in [Pil09b] some unconditional results including André's original one, but with an entirely different argument. In this same paper he also succeeded to mix the Manin-Mumford and André-Oort issues in special cases.

Moreover, in a more recent paper [Pil11] he goes much further and obtains a combination of the André-Oort, Manin-Mumford, and Lang's statements for arbitrary subvarieties of $X_1 \times \ldots \times X_n \times E_1 \times \ldots \times E_m \times \mathbb{G}_{\mathrm{m}}^{\ell}$ or of $X_1 \times \ldots \times X_n \times A$, where X_i are modular curves, E_i are elliptic curves, and A is an abelian variety, all defined over $\overline{\mathbb{Q}}$.

The *special points* are now the products of CM-points (relative to $\mathbb{C}^n$) times torsion points (relative to A or to $E_1 \times \cdots \times E_m \times \mathbb{G}_{\mathrm{m}}^{\ell}$). The *special subvarieties* are products of special subvarieties of the three main factors $X_1 \times \ldots \times X_n$, $E_1 \times \ldots \times E_m$ or $\mathbb{G}_{\mathrm{m}}^{\ell}$. As to each of these factors, we already mentioned the abelian (resp. toric) case: the special subvarieties are the torsion translates of abelian subvarieties (resp. algebraic subgroups). As to the "modular" factor $X_1 \times \ldots \times X_n$, the special varieties are, roughly speaking, defined by some CM constant coordinates plus several modular relations $\Phi_{n_{ij}}(c_i, c_j)$, relative to vectors $(c_1, \ldots, c_n) \in \mathbb{C}^n$. (If we view the modular curves as quotients of $\mathcal{H}$, the special subvarieties become images of points $(\tau_1, \ldots, \tau_n) \in \mathcal{H}^n$ with relations either of the shape $\tau_i = a_i$, a_i fixed quadratic imaginary numbers, or $\tau_r = g_{rs}\tau_s$, g_{rs} fixed in $GL_2^+(\mathbb{Q})$.)

These results of Pila are quite remarkable, especially taking into account that only a few unconditional cases of the André-Oort conjecture have been published after the one obtained by André.[10]

In Chapter 4 we shall sketch Pila's argument for André's theorem, in order to illustrate the method and to compare it with André's. This method of Pila is along the same lines of what we have seen in connection with Masser's questions; this time he relies on estimates not merely for rational points, but for *algebraic points of bounded degree* on transcendental varieties (he carries this out in [Pil09a] and [Pil11]). Also, the nature of these varieties is more complicated than before (they are not subanalytic), and in order to apply his results, Pila then relies on the fact that the varieties in question are of mixed exponential-subanalytic type.

Substantial effort also comes in characterizing the *algebraic* part of the relevant transcendental varieties, namely the union of connected semialgebraic positive dimensional arcs contained in it (we have already mentioned that these arcs have to be removed in order to obtain estimates of the sought type). This issue represents the geometric step of the method; it is the part taking into account the special subvarieties. This was a relatively easy matter in the paper [MZ08] but became more difficult in [PZ08] and [MZ10b] and is a major point in [Pil11].

We shall conclude the chapter by a brief discussion of Shimura varieties.

In the notes to this chapter we shall briefly discuss Edixhoven's approach to André's theorem, and then give a few definitions in the context of *definability* and *o-minimal structures*: these concepts concern collections of real varieties satisfying certain axiomatic properties, which are fundamental in the said analysis of rational points (by Pila and Pila-Wilkie).

[10] See, for instance, [Zha05] for some results of Shou-Wu Zhang on quaternion Shimura varieties.

A few words on effectivity. The mentioned method used jointly with Masser and separately with Pila for (relative) Manin-Mumford, and then by Pila toward the André-Oort conjecture, should be in principle effective in itself; but the tools on which it relies may or may not be effective, depending on the cases. For the Manin-Mumford statement and the problems of Masser, the necessary lower bounds for degrees are effective, and the upper bounds coming from the Bombieri-Pila-Wilkie papers should also be effective. (However, this has not yet been formally proved.)

On the contrary, the lower bound used in Pila's proofs concerning the André-Oort statement is not yet known to be effective: this is because it comes from Siegel's lower bounds for the class number of imaginary quadratic fields. (This leads only to an upper bound for the number of exceptions.)

Precisely the same kind of ineffectivity occurs in André's proof. However, inspection shows that one may in fact dispense with his (ingenious) opening argument, which is the only one using class-number estimates, at the cost of working out with additional precision the last part of his proof. (Here one also has to use effective lower bounds for linear forms in logarithms, in addition to effective lower bounds of Masser for algebraic approximations to values of the inverse of the modular function at algebraic points, a step which already appeared in the proof.) This variation leads to a completely effective statement. This is very recent work of L. Kühne [Küh11], and has been independently realized also in joint work with Yu. Bilu and Masser [BMZ11]; we shall give a sketch of this in Chapter 4, after reproducing André's proof.

Just to mention an explicit instance, we may think of the plane curve $x_1 + x_2 = 1$, which appeared already in the context of Theorem 1 as a motivating simple example; it might be sensible to take it into account again by asking about the CM-pairs on it:

Which are the pairs of CM-invariants j, j' such that $j + j' = 1$?

Now, $x_1 + x_2 = 1$ is not one of the $Y_0(n)$, so we know from André that the list of such pairs is finite. In fact, the said effective argument may be carried out explicitly in this case, and actually the special shape of the equation allows us to avoid all the ingredients from transcendental number theory which appear in the treatment of the general case. This leads to a simple elementary argument (carried out independently by Kühne and Masser) showing that

There are no such pairs.

For the sake of illustration, here are numerical values of some CM-invariants[11] (for which I thank David Masser), which certainly would somewhat suggest a priori that it is difficult to find solutions:

$$j(\sqrt{-1}) = 1728, \quad j(\frac{-1+\sqrt{-3}}{2}) = 0, \quad j(\sqrt{-2}) = 8000, \quad j(\sqrt{-5}) = 632000 + 282880\sqrt{5}.$$

A celebrated example (with which we see that $e^{\pi\sqrt{163}}$ is remarkably near to an integer) is also

$$j(\frac{-1+\sqrt{-163}}{2}) = -262537412640768000.$$

Less well known is (see [Mas03], p. 20)

$$j(\frac{-1+\sqrt{-427}}{2}) = -7805727756261891959906304000 - 999421027517377348595712000\sqrt{61}.$$

See also [Sil02], and further [Ser97], A.4, for the complete table of *rational integer* values of j at CM-points.

[11] See the end of Section 4.2 for some corresponding Weierstrass equations.

Appendixes

As announced in the preface, the book is concluded with seven short appendixes, the last six of which were written by David Masser. Let us briefly illustrate the corresponding contents.

Appendix **A**. This is concerned with the estimates for rational points on subanalytic surfaces, an ingredient at the basis of the method used in Chapter 3 (see the above paragraph "Upper bounds") and Chapter 4. We shall sketch a proof of a theorem of Pila, basic for the application to Masser's question in Chapter 3; this result (obtained in [Pil05]) concerns estimates for the number of rational points of denominator $\leq N$ on the "transcendental part" of a compact subanalytic surface. The method in part follows Bombieri-Pila's paper [BP89], which treated curves. Although we shall not touch the more general results of Pila needed for the applications of Chapter 4, we hope that this appendix may contribute to illustrate some of the main ideas in the context.

Appendix **B** (by D. Masser). Recall that Theorem 2 (in the above description of Chapter 1) predicts finiteness for the unlikely intersections of a curve in a torus; the proof yields an estimate for their number, which depends (among others) on the height of the curve. The question arose if this dependence can be eliminated. This appendix presents a deduction by Masser that it is indeed so, provided one assumes the Zilber's conjecture; for simplicity, Masser works with lines in $\mathbb{G}_{\mathrm{m}}^3$. (A generalization of these arguments to arbitrary varieties is also sketched in Subsection 1.3.8 below.)

Appendix **C** (by D. Masser). Consider an elliptic curve E with a finitely generated subgroup Γ of points, all defined over a rational field $K(t)$ (K being a number field); we assume that E has nonconstant invariant. This appendix presents a direct argument of Masser, proving an upper bound for the height of algebraic values t_0 of t for which the rank of the specialized group at t_0 decreases. This is a basic case of a general result by Silverman, containing the essentials of it. It is relevant here because a special case is in turn important in the finiteness proofs in Chapter 3 (and is carried out separately therein as Proposition 3.2).

Appendix **D** (by D. Masser). This presents a sketch of a proof of a lower bound by Masser for the degree d of (the field of definition of) a point of order n on an elliptic curve E defined over $\overline{\mathbb{Q}}$; this takes the shape $d \geq cn/\log n$, for an explicit positive $c = c(E)$. This is crucial for application to the proofs in Chapter 3. In that context one uses inequality (3.3.2), which essentially follows from the present bound after an explicit estimation of $c(E)$; this last step is indicated as well by Masser in this appendix.

Appendix **E** (by D. Masser). This appendix deals with transcendence of values of the modular function $j : \mathcal{H} \to \mathbb{C}$. Masser starts by proving the theorem of T. Schneider that if $j(\tau)$ is algebraic (for a $\tau \in \mathcal{H}$) then either τ is quadratic or transcendental. Masser also shows how to adapt this proof to a quantitative version in which the distance $|\tau - \beta|$ between τ and an algebraic β is bounded below effectively by $\exp(-c(1 + h(\beta)^k)$; this is important in the proof of André's theorem. (See Lemma 4.3, where this result is used with $k = 3 + \epsilon$, as in (4.3.6). This supplementary precision was obtained by Masser in [Mas75], but any k is sufficient for application to André's theorem.)

Appendix **F** (by D. Masser). We have mentioned the paper [BP89] by Bombieri-Pila on estimates for the distribution of rational points on real curves (discussed also in Appendix A). In this appendix, Masser gives a different argument, stemming from transcendence theory, to obtain similar bounds for the rational points lying on the graph of a transcendental real-analytic function.

Appendix **G** (by D. Masser). In this last appendix Masser considers a mixed Manin-Mumford-André-Oort statement, similarly to the end of Chapter 4; see especially Remark 4.4.4(ii), discussing Theorem 1.2 in [Pil09b]. Here Masser sketches a third argument for such a result, sticking for simplicity to a special case; he proves that there are only finitely many complex $\lambda = 0, 1$ for which the Legendre curve $y^2 = x(x-1)(x-\lambda)$ has complex multiplication and the point $P = (2, \sqrt{4-2\lambda})$ is torsion on it. This proof follows the method discussed in Chapter 3, and, contrary to Pila's, avoids recourse to Siegel's class-number estimates, as well as the appeal to the results of Pila [Pil09b] on quadratic points on definable varieties; in this direction it uses only Pila's results on rational points on subanalytic surfaces (sketched in Appendix A). This approach in particular eliminates the ineffectivity coming from Siegel's theorem.