



Torsion in algebraic groups and problems which arise

Umberto Zannier

Abstract. This article is based on the lecture that I had the honor and pleasure to deliver at the 8th European Congress of Mathematics in Portorož, Slovenia (originally planned for June 2020, then shifted to June 2021 for public health reasons). In the talk I tried to give an overview of some issues linked to torsion in algebraic groups, focusing on some recent research. Taking into account the purposes of reaching a large audience of mathematicians, from all subjects, I started with elementary general concepts, recalling some historical steps, before shifting to more specific themes which I was more familiar with. In these notes, I maintained the same principles, and only slightly expanded the contents of the lecture; indeed, I have not gone into any detailed argument.

1. Torsion in commutative algebraic groups

Torsion (etymology): the word *torsion* (in mathematics) often denotes a quantity, suitably defined in differential terms, which measures local “twisting” of a curve in Euclidean space (roughly speaking, it expresses “how far” the curve locally is from being a plane curve). However, in this exposition we shall adopt the usual *algebraic* meaning, namely according to the following definition.

Definition. An element g in a group G is *torsion* if $g^m = 1 = \text{identity of } G$, for some integer $m > 0$. (Such an m is called an *exponent* for g , whereas the minimal such m is called *the – exact – order* of g .)

This terminology apparently is not unrelated to the former one, as it seemingly originated from the structure of homology groups of spaces obtained by *twisting*. For instance, the real projective plane $\mathbb{P}_2(\mathbb{R})$, defined by gluing antipodal points in a closed half sphere, has the torsion group $\mathbb{Z}/2$ as its first homology group (over \mathbb{Z}).

A torsion element g as above generates a so-called *finite cyclic* group; now the etymology comes from the *circle*, because the powers g^n repeat *cyclically*: $\dots, g, g^2, \dots, g^{m+1} = g, g^{m+2} = g^2, \dots$ and generally $g^{n+m} = g^n, n = 0, 1, \dots$

2020 Mathematics Subject Classification. Primary 11-02; Secondary 11Gxx, 14Gxx, 14Kxx, 14Lxx.

Keywords. Number theory, Diophantine geometry, algebraic groups.

Indeed the *circle* comes into the picture beyond this simple intuition, through its topology (especially the fundamental group).

1.1. Algebraic groups

We shall consider some examples of torsion elements, and their structure, in *algebraic groups*: roughly speaking an algebraic group is defined first as an algebraic variety, i.e., a set of points satisfying a given system of algebraic equations in an affine or projective space, and then one has a group law expressed by polynomials in the coordinates.

An algebraic group is an irreducible variety if and only if it is connected, and, in general, it is anyway a finite union of translates of the connected component of the identity element (which is a normal subgroup).

In this article we shall meet only *commutative algebraic groups*, a property which entails that torsion elements form a subgroup.

For simplicity we shall consider only algebraic groups and points defined over the field \mathbb{C} of complex numbers and tacitly identify such a group with the set of its complex points. (However, this does not mean that we shall disregard the actual *minimal* field of definition of the points of interest for us, a field which may be small and is highly important for arithmetical information.)

Examples.

Additive algebraic group. The *additive algebraic group*, denoted by \mathbb{G}_a , is simply the affine line \mathbb{A}^1 as an algebraic variety. The group law is expressed additively by $(x, y) \mapsto x + y$. The set of complex points $\mathbb{G}_a(\mathbb{C})$ of \mathbb{G}_a is simply \mathbb{C} .

A torsion element $g \in \mathbb{C}$ of exponent m now satisfies $mg = 0$, hence $g = 0$, which means that there is no torsion other than 0 (as over any field of characteristic zero, whereas every element is torsion of exponent p in positive characteristic p).

Multiplicative algebraic group. The *multiplicative algebraic group* \mathbb{G}_m is the affine line deprived of the origin $\mathbb{A}^1 - \{0\}$ as an algebraic variety, with the algebraic group law $(x, y) \mapsto xy$. The set $\mathbb{G}_m(\mathbb{C})$ of its complex points is the multiplicative group of nonzero complex numbers $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$.

A torsion element $g \in \mathbb{C}^*$ of exponent m satisfies $g^m = 1$, so the torsion elements are precisely the (complex) roots of unity. There are m having exponent m ; these lie on the *unit circle* $S_1 := \{z \in \mathbb{C} : |z| = 1\}$, and they form the vertices of a regular m -gon in the complex plane \mathbb{C} .

Note the (analytic) exponential map $z \mapsto e^{2\pi iz}$, which sends homomorphically \mathbb{C} onto \mathbb{C}^* and has a kernel $\mathbb{Z} \cong \pi_1(\mathbb{C}^*)$; this is a non-divisible group, which explains torsion elements in the image (whereas there are no nontrivial ones in the domain).

Through this map we have the *analytic* isomorphism $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^* \cong \mathbb{C}/\mathbb{Z}$, the quotient of \mathbb{C} by a *discrete* subgroup (of rank 1).

Elliptic curves. The additive and multiplicative algebraic groups are curves (i.e., have dimension 1). However, they do not exhaust the possibilities for a curve to be a connected algebraic group. Indeed, (the) other fundamental examples are given by (complex) *elliptic curves*. They can be defined in the *projective* plane \mathbb{P}_2 by equations of the shape

$$E: y^2 = 4x^3 - g_2x - g_3 \text{ in } \mathbb{A}^2 + \text{point at infinity } O := (0 : 1 : 0) \text{ in } \mathbb{P}_2,$$

where $g_2, g_3 \in \mathbb{C}$ are such that $4x^3 - g_2x - g_3$ has no multiple roots, i.e.,

$$g_2^3 - 27g_3^2 \neq 0.$$

What is particularly remarkable is the existence of an algebraic-group (commutative) law among the points of each such curve. Namely, if we prescribe that the origin is the point at infinity O , then to add points $P, Q \in E$, we first draw the line through P, Q (or the tangent to E if $P = Q$) which (taking into account multiplicities) will intersect E in a third point R . The group law is such that $P + Q + R = O$, whereas $P + Q$ is the point opposite to R with respect to the x -axis.

This outstanding law (called classically the *chord and tangent process*, apparently observed first by Newton) indeed may be expressed by polynomials in the homogeneous coordinates and satisfies the group axioms (the associative law being not entirely trivial to check). It is fundamental in several respects, e.g., in the theory of Diophantine equations, since, when the curve has rational coefficients, it produces rational points out of rational ones.

Somewhat similarly to the case of \mathbb{G}_m , each of these curves is found to be *analytically* isomorphic to a (compact) *complex torus* \mathbb{C}/L , where L is again a suitable discrete subgroup, however now of maximal rank 2, i.e., a lattice in \mathbb{C} . The isomorphism occurs through the *Weierstrass exponential map*: $z \mapsto (\wp_L(z), \wp'_L(z))$, where \wp_L is the Weierstrass function associated to L :

$$\wp_L(z) = z^{-2} + \sum_{l \in L - \{0\}} ((z-l)^{-2} - l^{-2}).$$

This function is meromorphic on \mathbb{C} and admits L as its group of periods. (It sends L to O .)

The addition $+$ on \mathbb{C} (and on \mathbb{C}/L) then explains the group law on E in the sense that the former is transported to the latter by the said exponential.

In particular, it appears that now there are m^2 torsion elements of exponent m .

The complex elliptic curves, up to complex isomorphism, form a family of dimension 1 (parameterized by the so-called j -invariant $j(E) = 1728g_2^3/(g_2^3 - 27g_3^2)$,

which can assume any complex value). Together with \mathbb{G}_a and \mathbb{G}_m , they exhaust the isomorphism classes of complex (connected) algebraic groups of dimension 1.

Abelian varieties. *Abelian varieties* are the irreducible (or, equivalently, connected) projective algebraic groups. They are automatically commutative (the terminology “abelian” arising for a different reason).

Elliptic curves represent precisely the abelian varieties of dimension 1. But abelian varieties exist in any dimension: a simple example is a power E^r of an elliptic curve E , though this is extremely special. Other very important (though still special) abelian varieties arise as *Jacobians of (smooth) algebraic curves* of genus $g > 0$; such a Jacobian has dimension g .

Like for elliptic curves, every complex abelian variety, say of dimension g , is *analytically* isomorphic to a complex torus, i.e., a quotient \mathbb{C}^g/L where L is a (full) lattice; however for $g > 1$ not every complex torus is an abelian variety, a certain subtle “bilinear” condition on the lattice (existence of a Riemann *form*), in heavy part arithmetical, being necessary and sufficient.

Products. We may obtain other algebraic groups by taking products, e.g., of the form $\mathbb{G}_a^r \times \mathbb{G}_m^s$; the complex points are now vectors in \mathbb{C}^{r+s} , where the last s coordinates are nonzero and where the operations are coordinatewise (additive on the first r coordinates, multiplicative on the last s ones). For topological reasons the powers \mathbb{G}_m^s are sometimes called (*complex multiplicative*) *tori*.

Similarly, we may take products among the other algebraic groups we have seen. However, one should take into account that there exist *extensions* of algebraic groups, i.e., exact sequences $0 \rightarrow G_1 \rightarrow G \rightarrow G_2 \rightarrow 0$, where G is not (necessarily isomorphic to) the product $G_1 \times G_2$ (examples occur already in dimension 2, on taking $G_1 = \mathbb{G}_a$ or \mathbb{G}_m and $G_2 =$ an elliptic curve). When $G_1 = \mathbb{G}_m^s$ and G_2 is an abelian variety, any G in such an exact sequence is called a *semiabelian variety*.

1.2. Some results about torsion in algebraic groups

Additive case. We have already noted that \mathbb{G}_a has no nontrivial torsion in characteristic zero, thus in particular over \mathbb{C} .

Multiplicative case. In this case, we have recalled that the torsion elements of $\mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*$ are the complex roots of unity.

Through the exponential map $z \rightarrow e^{2\pi iz}$, the roots of unity correspond to $z =$ rational number, which raises a link with Number Theory.

Roots of unity naturally appear in describing *discrete periodical phenomena*. For instance one finds here *finite Fourier series*, i.e., linear combinations of exponential functions (on \mathbb{Z}) having roots of unity as bases; they are a discrete counterpart of the famous series expansions introduced systematically by Fourier, at the very heart of Harmonic Analysis. The finite Fourier series represent all periodic functions on \mathbb{Z}

and are of the utmost relevance in myriads of topics and applications, including Coding Theory, Combinatorics, Fast Multiplication, Group Theory, (Analytic) Number Theory, Numerical Analysis, and so on.

An effectivity issue. Let us pause to note that, already for roots of unity, even to *decide* whether a specific number (or, more generally, specific point of a given algebraic group) is torsion seems not completely obvious. For instance, consider the following:

$$\text{Small challenge: } \text{Is } \alpha := \frac{3 + 4\sqrt{-1}}{5} \text{ a root of unity?}$$

Note that α indeed lies in the unit circle S_1 , as $3^2 + 4^2 = 5^2$. So the natural test of computing $|\alpha|$ does not disprove the sought eventuality for this particular number.

Now, we can check whether $\alpha^2, \alpha^3, \dots$, or any *given* power α^m , is or is not equal to 1, but a possible torsion-exponent is not bounded *a priori*; so, unless we find 1 at some stage, we are left with an open possibility for the next check.

In conclusion, a little reflection may be needed to (see how to) answer such (type of) question(s) algorithmically in the general case. (Now a negative answer may be obtained using the Euler function value $\phi(q)$ for the degree over \mathbb{Q} of a root of unity of exact order q , which bounds the possible torsion order in terms of the degree of the given number. This works generally, but for the actual question maybe the simplest way is to observe that α is not an algebraic integer.¹ We invite the interested reader to seek several different arguments for answering the question.)

From a number theoretical viewpoint, Gauss (*Disquisitiones Arithmeticae 1801*) was the first to study in depth the arithmetical properties of roots of unity. In particular, this led him to criteria for *constructing a regular n -gon with ruler and compass* (ancient problem of Greek mathematics). For instance this is possible for

$n = 3, 4, 5, 6, 8, 10, 12, 15, 16, 17, \dots$, but not for $n = 7, 9, 11, 13, 14, 18, 19, \dots$

As is well known, Fermat primes $2^{2^k} + 1$ play a heavy role here. . . .

In fact, already a few years before the publication of the *Disquisitiones*, Gauss had succeeded to construct the regular polygon of 17 sides, obtaining in practice the remarkable equality

$$16 \cos \frac{2\pi}{17} = -1 + \sqrt{17} + \sqrt{34 - 2\sqrt{17}} + 2\sqrt{17 + 3\sqrt{17} - \sqrt{170 + 38\sqrt{17}}}.$$

We may say that Gauss anticipated the *Galois theory of the cyclotomic fields*; in fact, in particular he defined the so-called *Gaussian periods*, which *a posteriori* turn

¹This fits with a well-known theorem of Kronecker: “Roots of unity are those algebraic integers having all conjugates of complex absolute value 1”, which may be rephrased as: “An algebraic number is a root of unity if and only if *all* its absolute values are 1”.

out to be suitable invariants for subgroups of automorphisms. (They may be also conceived as values of certain *finite Fourier series* alluded to above.) For instance, Gauss obtained from them “explicit” generators for all the subfields of a *principal* cyclotomic field $\mathbb{Q}(e^{2\pi i/p})$, where p is a prime number. (They are highly important for other reasons as well.) In particular, Gauss expressed, through the famous *Gauss sums*, any number \sqrt{n} , $n \in \mathbb{Z}$, as a sum of roots of unity, which is not at all obvious. This also started the theory of abelian extensions of \mathbb{Q} and of number fields (so-called “Class-field theory”).

In this case of the roots of unity (through viewpoints introduced by Deuring, . . . , Tate, . . . , Grothendieck, . . .) the Galois groups which arise may be seen as an algebraic manifestation and realization of the *monodromy (group) of the circle* $S_1 = \{z \in \mathbb{C} : |z| = 1\}$. For instance, we have

$$\pi_1(S_1) = \pi_1(\mathbb{C}^*) = \mathbb{Z},$$

and its finite quotients are the $\mathbb{Z}/(m)$ which correspond to the finite covers of S_1 , and so to the homomorphisms

$$S_1 \rightarrow S_1: z \mapsto z^m,$$

with kernel the group $U_m = e^{2\pi i\mathbb{Z}/m} \cong \mathbb{Z}/(m)$ of m th roots of unity. So U_m is the topological covering group and the Galois group over \mathbb{Q} acts on it, and we have

$$\text{Aut}(U_m) \cong (\mathbb{Z}/(m))^* \cong \text{Gal}(\mathbb{Q}(U_m)/\mathbb{Q}) = \text{Gal}(\mathbb{Q}(e^{2\pi i/m})/\mathbb{Q}),$$

as proved essentially by Gauss. So the algebraic Galois group of the corresponding field extension equals the (abstract) automorphism group of the topological covering group.

Elliptic case. Now the theory of torsion elements is again highly interesting, rich, and actually (much) more difficult than in the cyclotomic case. We have already recalled that there are m^2 elements of exponent m . The coordinates of these points generate (over the ground field $\mathbb{Q}(g_2, g_3)$) a field which is found to contain the cyclotomic field $\mathbb{Q}(e^{2\pi i/m})$, so we may say that the cyclotomic case recalled above falls just as a special piece of the elliptic theory.

The torsion points now lie on a space which may be identified with the product $S_1 \times S_1$ of two circles (a torus), and the topological covering group corresponding to torsion points of order m is now $(\mathbb{Z}/(m))^2$. The elements of the Galois group again correspond to automorphisms of the covering group and thus may be viewed inside the finite matrix group $\text{GL}_2(\mathbb{Z}/(m))$. A fundamental issue is to understand the image of the Galois group (as m varies). This Galois theory somewhat depends on the coefficients g_2, g_3 .

The “generic” case of transcendental j -invariant had been dealt with by Fricke & Weber between the XIX and XX centuries: they proved that the image is essentially the “largest possible one” (i.e., $\mathrm{SL}_2(\mathbb{Z}/(m))$) if we work over \mathbb{C}).

The algebraic case lies much deeper, and suppose to fix ideas that $g_2, g_3 \in \mathbb{Q}$. There are two essentially different subcases, according to whether the ring of endomorphisms of the elliptic curve is “trivial” (i.e., \mathbb{Z}) or not; the latter case, called *Complex Multiplication*, is “exceptional” in various ways (now the endomorphism ring is an *order* in some imaginary quadratic field).

Already Gauss (beginning of Chapter VII of *Disquisitiones*) foresaw the interest and depth of this issue in some of these situations, which he interpreted as analogous to cyclotomy, i.e., as the (arithmetical) theory of the dissection of a *lemniscate* (in place of a circle) into equal parts. It is also interesting that the general case of prime m had been considered by Galois (in a letter to Chevalier, 29 May 1832), especially from the viewpoint of *solvability by radicals* of the corresponding algebraic equations.

We skip any other detail and only recall a few basic more modern achievements.

Some elliptic results

- A deep landmark result on the above-mentioned Galois image is Serre’s *Open Image Theorem* (70s): in a sense it extends Gauss’ achievements (and more) to the most general elliptic case, proving that the Galois image is as large as possible (compatibly with the endomorphism ring) up to bounded index. (We omit a precise statement, which would lead us outside the scope of these notes.)
- Another very important and deep theorem is due to Mazur (70s), who proved that for $g_2, g_3 \in \mathbb{Q}$ the possible torsion orders of *rational* torsion points never exceed 12. This result corresponds to finding all rational points on suitable *modular curves*, providing a link of the present topic with major questions in the theory of Diophantine equations.
- Merel 1994, with some new ingredient, extended this kind of result to number fields other than \mathbb{Q} (some independent work being due to Kamienny & Parent, and previously to Demianenko & Manin in the case of prime-power torsion order).

Case of abelian varieties. The arithmetic and Galois structure of torsion elements is even more difficult than the special elliptic case. But nowadays there has been great progress, thanks to the work of Deligne, Bogomolov, Faltings, Serre, . . . , Masser & Wüstholz, . . . , Mazur, Ribet, Pink, Tamagawa, Cadoret, . . .

2. Algebraic relations among torsion points

We have recalled some results on individual torsion points. Let us now see some problems on *relations* among torsion points.

An old significant example comes from Gordan 1877 who studied the equation

$$\cos 2\pi x + \cos 2\pi y + \cos 2\pi z + 1 = 0, \quad x, y, z \in \mathbb{Q},$$

with the purpose of classifying the finite subgroups of $\mathrm{PGL}_2(\mathbb{C})$. On writing

$$2 \cos 2\pi z = e^{2\pi iz} + e^{-2\pi iz},$$

we see that this amounts to a certain algebraic equation among the three roots of unity $e^{2\pi ix}$, $e^{2\pi iy}$, $e^{2\pi iz}$, or else a certain (inhomogeneous) linear relation among three roots of unity and their reciprocals.

Later on, general *linear* equations in roots of unity were studied systematically, in particular by Mann 1965 and Conway & Jones 1976, in the setting of what the latter authors called *trigonometric diophantine equations*. These results in particular bounded the maximal torsion order in a linear equation with nonzero constant term and no vanishing subsums (with coefficients in \mathbb{Q}). As a very special instance, their conclusions very easily imply that

the only triangles with rational sides and angles rational multiples of π are equilateral,

and similar results follow for polygons with a given number of sides.

More recent applications appear, for example, in the work of Gross, Hironaka, and McMullen (to cyclotomic factors of E_n -Coxeter polynomials, 2009), of Bourgain, Gamburd, and Sarnak (to Markov surfaces 2016), of Kedlaya, Kolpakov, Poonen, and Rubinstein (to rational angles among vectors in \mathbb{R}^3 , 2020), and in a joint work of the author with Dvornicich & Veneziano (to rational angles in plane lattices, 2020).

Uniform *quantitative* results (regarding the number of solutions of a given linear equation) were proved, e.g., by Schlickewei, Evertse, Beukers & Smyth, and in a joint work of the author with Bombieri, also towards the conjecture of Lang to be discussed in the next section. (Subsequently these results have been quantitatively refined by several authors, including Amoroso & Viada and Martinez.)

2.1. The conjecture of Lang

Independently of the above authors, Lang had raised in the 60s the related problem of studying polynomial equations:

$$F(\theta, \zeta) = 0, \quad \theta, \zeta \text{ roots of unity, of } \textit{unrestricted} \text{ exponent.}$$

Note that such a pair (θ, ζ) is a *torsion point* on the plane curve $F(x, y) = 0$, viewed inside \mathbb{G}_m^2 .

Let F be given. As expected by Lang, there can be infinitely many solutions of the said shape only if F has a binomial factor of the shape $ax^m + by^n$ or $ax^m y^n + b$; this

was quickly proved by Ihara, Serre, and Tate (and proofs may be got also through the previously mentioned results on linear equations, on considering monomial terms).

This was later extended to arbitrary dimensions by M. Laurent and Sarnak & Adams, proving (among other things) the *conjecture of Lang*:

An algebraic subvariety of \mathbb{G}_m^r can have infinitely many torsion points only if it contains a positive dimensional special subvariety, i.e., a translate of an algebraic subgroup by a torsion point.

More precisely, their results yield the following theorem.

Theorem 2.1. *Let Σ be any set of torsion points inside \mathbb{G}_m^r . The Zariski closure of Σ is a finite union of translates (by torsion points) of algebraic subgroups.*

It is moreover not difficult to see that any connected algebraic subgroup of \mathbb{G}_m^r can be defined by finitely many equations of the shape $x_1^{a_1} \cdots x_r^{a_r} = 1$ and is (algebraically) isomorphic to some \mathbb{G}_m^h ($h \leq r$). Hence in practice the principle is that

every prescribed algebraic relation within varying torsion elements can be explained in finite terms by a multiplicative structure of algebraic group.

Methods. The Galois theory of Gauss is a precious tool in all these achievements, though also other ingredients are relevant.

As for the previously mentioned work, later this had several applications.

For instance we quote the work by Sarnak (on *Betti numbers of congruence groups*, 1994), by Ailon & Rudnick (on $\gcd(f(t)^n - 1, g(t)^n - 1)$, 2004), by Kurasov & Sarnak (on *crystalline measures*, 2020).

2.2. Multiplicative relations on curves – unlikely intersections

The mentioned issues on torsion points may be extended to deal with more general *multiplicative relations* among coordinates of points on a curve X inside a torus \mathbb{G}_m^r . That is, we weaken the condition that all the coordinates are torsion and only impose that a certain number of independent multiplicative relations hold among the coordinates.

It is easy to see that if we prescribe on the irreducible curve X a single multiplicative relation, i.e., of the shape $x_1^{a_1} \cdots x_r^{a_r} = 1$ with $(x_1, \dots, x_r) \in X$, then we obtain infinitely many points as (a_1, \dots, a_r) varies through all nonzero integer vectors; this corresponds to intersect X with the union of all proper algebraic subgroups of \mathbb{G}_m^r . However, it turns out that imposing another such relation, independent with the former but otherwise arbitrary (which corresponds to intersect X with the union of algebraic subgroups of codimension ≥ 2), yields only finitely many points, unless the curve X is *special* in the sense that it is contained in a *proper* algebraic subgroup

of \mathbb{G}_m^r . (For $r = 2$ we obtain nothing new since imposing two independent relations yields torsion coordinates, but for $r > 2$ this is a weaker restriction, making the theorem stronger.)

With the more demanding assumption that X is not contained in any proper *translate* of algebraic subgroup, this was dealt with in a joint work of the author with Bombieri & Masser 1999, and later proved in the sharper form by Maurin 2008, relying partly on methods by Rémond. (This case is more difficult; for instance it contains implicitly the so-called Mordell–Lang context for tori.) A different approach for this stronger theorem was found later by the former authors with Habegger 2010 (this time *using* the results of Mordell–Lang type). A still further approach with the stronger assumption appeared in a joint work of the author with Capuano, Masser, and Pila 2016, based on the counting method alluded to below; this argument has the advantage of extending to the abelian context (but not containing the sharper form).

Several results followed by other authors as well, also for some higher-dimensional analogues, and further in the abelian case.

The topic, sometimes called *Unlikely Intersections*, was independently raised also by Zilber 2002 (with entirely independent motivations from Logic) and again independently by Pink 2005.² They put forward certain general conjectures still widely open (those of Pink embracing still further realms). These conjectures dealt also with abelian varieties in place of tori, where exact analogues may be stated. We shall briefly discuss this context in the next subsection.

2.3. The conjecture of Manin–Mumford

A motivation for the above-mentioned problems stated by Lang had been a conjecture formulated independently by Manin & Mumford in the 60s.

Manin–Mumford conjecture. *A curve of genus ≥ 2 embedded in its Jacobian variety has only finitely many torsion points.*

This may be indeed seen as an analogue (of more difficult nature) for abelian varieties of some of the above problems for multiplicative tori. It became a theorem due to Raynaud in the 80s; he was able to prove, more generally, the analogue of Lang’s conjecture above, and for arbitrary abelian varieties (not merely Jacobians) and arbitrary subvarieties. Several other proofs then followed, due, e.g., to Serre, Coleman, Hindry, Buium, Hrushovski, Pink & Roessler, M. Baker & Ribet.

Still other proofs (by Bilu 1997 for \mathbb{G}_m^r and Szpiro, Ullmo, and S. Zhang 1997 for the abelian case) gave stronger results of *Galois equidistribution* of the conjugates of torsion points when the degree of the field of definition of the points grows. Moreover,

²Certain rather special cases had been raised earlier by Schinzel, with still different language and motivations, coming mainly from his theory of reducibility of lacunary polynomials.

these proofs worked also for points of “small height”.³ Remarkable *uniform* estimates here (e.g., for the number of torsion points on the curve) have been given very recently by Kuehne 2021.

A further proof was found in a joint work of Pila and the author (2009): this relied on the analytic isomorphism of a complex abelian variety with a complex torus, in which torsion points correspond to rational points (as in the case of roots of unity). Then one reduces to *counting rational points on suitable analytic subvarieties* of the torus and comparing bounds from below (coming from the large degree of torsion points – work by Masser) and from above. This last step is done through estimates by Bombieri & Pila 1989, Pila, and finally Pila–Wilkie 2006. In turn, this last work involves the (model)-theory of the so-called *o-minimal structures* (developed by van der Dries et al.).

2.4. “Special points” and the André–Oort conjecture

As far-reaching analogues of torsion points, one may consider the so-called *special points* in *Shimura varieties*. An important kind of such varieties arises as moduli spaces of abelian varieties with certain properties (i.e., parametrizing abelian varieties of given dimension with supplementary symmetries). The *special points*, playing the role of torsion points, are those corresponding to Complex-Multiplication abelian varieties. Moreover, one may also define *special subvarieties* of positive dimension, analogues of the translates (by torsion points) of algebraic subgroups (in the conjecture of Lang) or of abelian subvarieties (in the theorem of Raynaud, in turn analogue of Lang’s for abelian varieties).

We skip any formal definition, since the context is quite technical, but we note that one may formulate statements analogue to the above ones. A very relevant instance is the *André–Oort conjecture*, raised by André 1989 and Oort 1990s independently. Once that the above terminology has been introduced in precise terms, a possible phrasing of it is as follows:

The Zariski closure of a set of special points is a finite union of special subvarieties.

This formulation reminds of what we have seen in the multiplicative and abelian cases.

After the proof of a special case by André (i.e., the significant case of CM-points on a curve in the plane \mathbb{A}^2 , viewed as representing pairs of elliptic curves), the above-

³The *height* of a point with algebraic coordinates is a real non-negative number which measures its arithmetical complexity; one may define a *canonical* height on the algebraic points of a commutative algebraic group, which vanishes precisely at torsion points; we cannot pause here on this concept, introduced first by Weil, despite its great relevance.

mentioned *counting method* was applied by Pila to this context, proving substantially more general instances.

A final step for the full conjecture (for the moduli space \mathcal{A}_g) was finally devised by Tsimerman 2015 after many important intermediate results and steps, in particular by Colmez, Edixhoven, Gao, Klingler, Pila, Pila & Tsimerman, Ullmo & Yafaev, Yuan & S. Zhang, . . . (A still more general form of the conjecture has been obtained very recently by Pila, Shankar, and Tsimerman, relying also on further ingredients provided by Binyamini and by Esnault & Groechening.)

Several other results in a similar spirit have been obtained and much work in the context is still in progress.

Dynamical analogues. Still further analogues of special points occur in *dynamics*, which we describe roughly as the study of iterates $f, f^{\circ 2} := f \circ f, \dots, f^{\circ n}, \dots$ of a map $f : X \rightarrow X$ from a space X to itself. The simplest examples (already leading often to very difficult problems) occur with rational maps $f : \mathbb{P}_1 \rightarrow \mathbb{P}_1$. As possible analogues of torsion points one can consider *preperiodic points* for f , i.e., the points $x \in X$ such that $f^{\circ n}(x) = f^{\circ m}(x)$ for some integers $n > m$ (so that the sequence $(f^{\circ r}(x))_{r \in \mathbb{N}}$ is finite). For instance, if $X = \mathbb{G}_m$, $f(x) = x^d$ (any $d \geq 2$), the preperiodic points are precisely the torsion points. One may formulate analogues of the above statements, and some quite nontrivial remarkable results have been proved, mainly due to the work of M. Baker, Bell, DeMarco, Ghioca, Hsia, Mavraki, Scanlon, Silverman, Szpiro, Tucker, Yuan, S. Zhang, . . . , among others. However, only a partial picture has been obtained to date in this direction compared to the original context of torsion points, and even a satisfactory formulation of suitable *complete* conjectures seems not to have been reached so far.

3. Torsion in families of algebraic groups

We have briefly discussed torsion in individual algebraic groups, and algebraic relations among them. To go one step further, we can consider torsion conditions in algebraic groups (and points) varying in families. The multiplicative group \mathbb{G}_m does not admit genuine “variation”, but already for elliptic curves we have truly *nonconstant families*. A typical and historically relevant instance of this is the *Legendre family* of elliptic curves, defined by

$$\mathcal{L}_\lambda: y^2 = x(x-1)(x-\lambda) + \text{point at infinity } O,$$

where λ is a complex parameter in $\mathbb{C} - \{0, 1\}$. For each $b \in \mathbb{C} - \{0, 1\}$ up to two exceptions, there are only six values of λ producing a curve isomorphic to \mathcal{L}_b , and each complex elliptic curve is isomorphic to some \mathcal{L}_b , so the family indeed is intrinsically not “constant”.

Regarding families of points (also called *sections*), we may consider, just as a simple instance, the points

$$P_\lambda = (2, \sqrt{2(2-\lambda)}) \in \mathcal{L}_\lambda,$$

where the choice of the sign is immaterial for us.

It may be shown that

- (1) P_λ is not identically torsion on \mathcal{L}_λ (i.e., there is no integer $m > 0$ such that $mP_\lambda = O$ for all λ), but
- (2) P_b becomes torsion on \mathcal{L}_b (of unrestricted exponent) for an infinite, even dense, set of $b \in \mathbb{C}$. This set consists of algebraic numbers, and the corresponding minimal torsion exponents tend to ∞ ;
- (3) these numbers b have *bounded height*. So for instance *there are only finitely many rational or even quadratic irrational ones*, and in fact the degree over \mathbb{Q} of these numbers tends to ∞ . (Néron had previously shown that they form a so-called *thin* set in any given number field.)

Property (1) follows from the general principle that torsion points are *unramified* except above the locus of *bad reduction*. Property (2) may be proved through the Betti map, mentioned below. Property (3) follows from results by Silverman & Tate 1980s. Properties (2) and (3) actually hold for all sections (defined over $\overline{\mathbb{Q}}$) satisfying (1).

Further Galois-equidistribution results for these numbers b are due to DeMarco & Mavraki 2019. Note that the equidistribution here does not concern the (conjugates of the) hypothetical torsion points, but regards the (conjugates of the) values b for which P_b is a torsion point. Hence this result has a quite different meaning with respect to the previously mentioned equidistribution theorems of Bilu and Szpiro, Ullmo, and Zhang. This equidistribution implies in particular the above-mentioned complex density.

For the actual choice of family (using the Betti map appearing below) one can also prove density of the relevant b in the real half-line $(-\infty, 2)$, so that $P_b \in \mathcal{L}_b(\mathbb{R})$. On the other hand, a joint work of the author with Lawrence observes that we never have p -adic density for this set.

3.1. Masser's problem and the Pink conjectures

Masser considered a second family of points, for instance

$$Q_\lambda = (\lambda + 1, \sqrt{\lambda(\lambda + 1)}) \in \mathcal{L}_\lambda.$$

The same remarks (1), (2), and (3) hold for this family, and moreover P_λ, Q_λ may be shown to be *generically linearly independent* on \mathcal{L}_λ , i.e., if $rP_\lambda + sQ_\lambda = O$ for certain integers r, s and all λ , then $r = s = 0$.

From (3) we see that the values b of λ for which each point becomes separately torsion form a *sparse* set, so Masser asked the following.

Masser’s question. Is the “*doubly sparse*” set

$$\{b \in \mathbb{C} : P_b, Q_b \text{ are both torsion on } \mathcal{L}_b\}$$

even a finite set?

Here Galois groups of torsion points do not give enough information, essentially because the degree of the relevant numbers “ b ” is unbounded (and actually tends to ∞).

Using the above-mentioned counting method (and other tools), Masser & the author (2008) gave an affirmative answer to the question, actually to its natural generalization to other pairs of families and sections.

Later this was further extended to arbitrary algebraic pencils of abelian varieties and in other directions (e.g., of *Unlikely Intersections* type), also by M. Baker, Barroero, Bertrand, Capuano, Daw, DeMarco, Dill, Habegger, G. Jones, Orr, Pila, Pillay, H. Schmidt, Stoll, Tsimerman,

Some of these results may be seen as *relative* analogues of the Manin–Mumford conjecture (i.e., where the ambient abelian variety moves in a family), and some other ones as *dynamical* analogues (i.e., when the torsion points are replaced by *preperiodic points* with respect to suitable rational maps).

The problem of Masser was recognized as a special case of conjectures by Pink (and also of Zilber in other cases). As alluded above, these conjectures deal with much more general contexts (including the André–Oort one) and are still widely open.

3.2. The Betti map

The counting method alluded to above worked for families and points defined over $\overline{\mathbb{Q}}$, but some of the tools failed over \mathbb{C} . This obstacle was overcome in a joint work of the author with Corvaja & Masser 2017 by *specialization*, to reduce to the algebraic case.

This gave as a byproduct somewhat analogous conclusions for families parameterized by spaces of dimension > 1 .

Specialization appeared delicate because of certain possible degeneracies, difficult to exclude *a priori*. To get control on this, a relevant tool came from the so-called (real analytic) *Betti map*: it gives the real coordinates of the point, in terms of a lattice basis for the torus representing the abelian variety, the basis varying locally holomorphically in the family.

Example 3.1. In the case of the Legendre family, consider a lattice $L_\lambda \subset \mathbb{C}$ such that $\mathbb{C}/L_\lambda \cong \mathcal{L}_\lambda$ (for instance through a Weierstrass exponential giving the Legendre equation). Then, e.g., in the region $\mathcal{R} \subset \mathbb{C}$ defined by $\max(|\lambda|, |1 - \lambda|) < 1$, by

formulae going back to the XIX century, one can express a \mathbb{Z} -basis of L_λ in terms of hypergeometric functions, in fact as $L_\lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, where $\omega_1 = i\pi F(1-\lambda)$, $\omega_2 = \pi F(\lambda)$ and where $F(\lambda) = \sum \binom{-1/2}{n}^2 \lambda^n$. For a given $\lambda \in \mathcal{R}$ and a point $Q \in \mathcal{L}_\lambda$ we may take a representative for Q in \mathbb{C}/L_λ of the shape $\beta_1\omega_1 + \beta_2\omega_2$ with $\beta_1, \beta_2 \in \mathbb{R}/\mathbb{Z}$. Then by definition the β_i are the Betti coordinates of Q and the Betti map takes the value (β_1, β_2) at Q .⁴

The Betti map is highly relevant in our context because its *rational values correspond precisely to torsion points*. We have already mentioned some proofs where essential use is made of this map.

The Betti map appeared implicitly already in a work by Manin 1960s and was recently studied (for higher dimensions) in a work of Voisin 2019 and of André, Corvaja & the author 2020, with further contributions by Gao and applications by Voisin and by Dimitrov–Gao–Habegger and Kuehne.

4. Some applications

4.1. Pell equations in polynomials

The *Pell equation*

$$x^2 - y^2 D = 1, \quad D \text{ non-square positive integer,}$$

to be solved in *integers* $x, y \neq 0$, is a celebrated Diophantine equation, proposed in fact by Fermat in the XVII century but actually having roots in ancient mathematics. It is linked with many important issues in Number Theory, such as integral points on curves (especially general affine conics), class-numbers and units of quadratic rings, orthogonal groups over \mathbb{Z} , Diophantine approximation and continued fractions, and so on.

There is also a *polynomial* analogue, more recent and apparently less known, but in fact also old, studied for instance already by Abel 1826, where $D = D(t)$ is a (complex, for instance) polynomial of even degree $2d$ and not a square, and one seeks *polynomial* solutions $x(t), y(t) \neq 0$. Following a suggestion of Serre, this equation may then be called *Pell–Abel equation*.

As in the classical case, a possible *nontrivial* solution generates infinitely many ones through the formulae $x_n \pm y_n \sqrt{D} = (x \pm y \sqrt{D})^n$, $n \in \mathbb{Z}$ (and all solutions are generated in this way, up to sign, from a “minimal” one).

⁴This map may be defined in any given open simply connected region, like the above \mathcal{R} , and we can cover the domain $\mathbb{C} \setminus \{0, 1\}$ with such regions. Then the map depends on a choice of basis for a given region and is subject to monodromy as we travel through loops meeting several regions.

For the Pell–Abel equation, when $\deg D = 2$ there are always solutions (over \mathbb{C}), and the polynomials $x(t)$, $y(t)$ which arise are related to Chebyshev polynomials. But if $\deg D \geq 4$, contrary to the classical case, it is generally unexpected to have solutions (unless we work over a finite field). This assertion can be put on a rigorous ground for instance using the Betti map. Indeed, the issue is linked with *torsion* in the Jacobian of the smooth complete hyperelliptic curve $H = H_D$ defined affinely by $u^2 = D(t)$. In fact, denoting ∞_{\pm} the poles of t on H , it is not difficult to prove that *(nontrivial) solutions exist if and only if the class of the divisor $\infty_+ - \infty_- \in \text{div}(H)$ has finite order in the divisor class group – i.e., in the Jacobian – of H .*⁵ Abel gave a translation of such condition in terms of the continued fraction for $\sqrt{D(t)}$ being periodic (as happens in the classical case).

The polynomials $D(t)$ for which nontrivial solutions of the Pell–Abel equation exist are sometimes called *Pellian*.

In a joint work of the author with Masser we studied some 1-parameter families for fixed d , like $D_{\lambda}(t) = D_{d,\lambda}(t) := t^{2d} + t + \lambda$. As said, for $d = 1$, D_b is always Pellian. For $d = 2$ we easily realized that $D_b(t)$ is Pellian for infinitely many $b \in \mathbb{C}$ (satisfying (3) of Section 3), whereas we proved that for $d = 3$ there are only finitely many such values. We then extended the analysis to arbitrary 1-dimensional families of polynomials $D(t)$ of higher degree and those results would lead, for example, to the following theorem.

Theorem 4.1. *For any $d \geq 3$, there are only finitely many $b \in \mathbb{C}$ for which the Pell–Abel equation for $D_{d,b}(t)$ is solvable.*

We note that 0 lies in all these sets $\mathcal{P}_d := \{b \in \mathbb{C} : D_{d,b} \text{ is Pellian}\}$, for we have

$$(2t^{2d-1} + 1)^2 - (2t^{d-1})^2 D_{d,0}(t) = 1.$$

Open question. Is the union $\bigcup_{d \geq 3} \mathcal{P}_d$ of these finite sets itself finite?

If the answer is at all affirmative, it appears to require new tools to be proved, since the method that we used to deal with each single degree $d \geq 3$ is not completely uniform as d varies.

The Pell–Abel equation, similarly to the original version, appears in many mathematical topics; just to mention a recent instance, it has been studied by Kollar in connection with decidability issues and the Hilbert X problem over function fields. (We recall that the usual Pell equation had been used by Matijasevic in his final step solving the original Hilbert X problem.)

⁵A *generalized Jacobian* has to be considered if $D(t)$ has multiple roots. This link with Jacobians may be viewed as somewhat analogue of Dirichlet class number formula for real quadratic fields, the analogy being closer if we work over finite fields.

4.2. Integration in finite terms

The problem of expressing indefinite integrals in terms of “simple” functions goes back to long ago and appeared among the first examples and motivations for *differential algebra*. In this direction, we recall for instance the following (more or less classical) definition.

Definition. We call *Integrable in Finite Terms* (abbr. IFT) a differential whose (indefinite) integral can be expressed by a finite tower of operations either of algebraic type, or by taking exponentials or by taking logarithms (starting from rational functions). We also call *elementary* an integral which can be likewise expressed.

Even recently, much attention has been given to the study of possible algebraic relations among (definite) integrals of algebraic functions, special cases of *periods* (after Grothendieck, . . . , Kontsevich & Zagier, . . .), a topic not entirely unrelated with this theme.

We have already mentioned Abel in connection with Pell’s equations in polynomials, and indeed his research involved also elementary integration. Subsequently the matter was studied by authors like Chebyshev, Liouville, Ritt, Kolchin, . . . , giving rise for instance to Differential Galois theory.

More recently, J. Davenport investigated *pencils* of algebraic differentials, to be integrated in finite terms; he sought to understand *whether*,

if the general member of the family cannot be likewise integrated, the same happens for the special members, up to finitely many exceptions.

Together with Masser we found how to establish when this type of assertion is correct, and we also found some counterexamples.

By means of a criterion of Risch and other considerations, it turns out that the analysis for such results in fact involves *torsion*, now in *generalized Jacobians*, which are algebraic groups obtained as extensions of usual Jacobians by products of groups of type \mathbb{G}_a or \mathbb{G}_m .

Jointly with Masser, we carried out this, applying in particular some of the above results, and here are special cases of the output (all results joint with Masser 2018–2020).

Theorem 4.2. *There are only finitely many $b \in \mathbb{C}$ such that the integral $\int \frac{(2z+b) dz}{\sqrt{z^4+z+b}}$ is elementary.*

Example 4.3. The special value $b = 1/2$ is in the said finite set:

$$\int \frac{(2z + 1/2) dz}{\sqrt{z^4 + z + 1/2}} = \frac{1}{2} \log(4z^4 - 4z^3 + 2z^2 + 2z - 1 + (4z^2 - 4z + 2)\sqrt{z^4 + z + 1/2}).$$

This corresponds to a torsion point of order 4 in an extension by \mathbb{G}_a of the elliptic curve $w^2 = z^4 + z + (1/2)$.

The next example yields a negative answer towards Davenport’s issue.

Example 4.4 (Counterexample). The differential $\frac{z dz}{(z^2-t^2)\sqrt{z^3-z}}$ (over $\mathbb{C}(t)$) is not identically IFT but it becomes IFT for infinitely many specializations $t \rightarrow b$.

In this example, note the underlying elliptic curve $w^2 = z^3 - z$ with CM: this is no coincidence, since it can be shown that if the (usual) Jacobian of the underlying curve (corresponding to the differential) does not contain CM elliptic curves, then Davenport’s expectation is correct.

4.3. Elliptical billiards

Further applications of some of the results are to *elliptical billiards*, namely billiard tables whose border is an ellipse and such that consecutive segments of billiard trajectory obey the usual law of reflection at the border.

Work going back to Poncelet and Jacobi shows that to such a billiard one can associate an elliptic family. In fact, it may be shown by nice arguments of Geometry, of type almost going back to Euclid, that all segments in a *given* billiard trajectory are tangent to a same conic, confocal with the ellipse, the so-called *caustic*. This caustic varies in a family of dimension 1. If the caustic is given, then the set of pairs (P, l) , where P lies on the ellipse and l is a line through P tangent to the caustic, describes a curve of genus 1 embedded in \mathbb{P}_1^2 . This curve becomes an elliptic curve after choice of an origin, whence, as the caustic varies, we obtain the said elliptic family.

A choice of a slope for a billiard shot from a given point yields a section of this family (depending on the point and parameterized by the slope). The *torsion* values of such a section correspond to the trajectories which are *periodic*, whose analysis is a main issue in the study of billiards.⁶

In this frame, on applying some of the above-mentioned results, in a recent joint work with Corvaja (2021) we deduced certain finiteness theorems for periodic trajectories in such billiards. For instance, we have the following conclusion.

Theorem 4.5. *For each $\alpha \in (0, \pi)$ there are only finitely many periodic pairs of billiard shots from a given point in an elliptical billiard such that the initial directions form an angle α .*

This may be shown to be not generally true for rectangular billiards.

⁶Part of this is a special case of a famous theorem of Poncelet, dealing with more general pairs of conics. The context has been generalized to higher dimensions by Griffiths & Harris 1977, which raises again questions related to the present realm.

Another finiteness conclusion (proved however with results of “Unlikely Intersections” type going beyond torsion – see above) concerns the set $T_{P,Q,R}$ of billiard trajectories which connect two given points P, Q and pass through another given point R : for instance, we have the following theorem.

Theorem 4.6. *If Q is a hole (i.e., lies on the boundary) and P, R are not both foci of the ellipse, then the set $T_{P,Q,R}$ is finite.*

It is somewhat curious that some of these results in the degenerate case of a circular billiard are related to the above discussion around Lang’s conjecture.

Still further conclusions in the same spirit may be stated, e.g., concerning *boomerang* billiard shots. The link with the algebraic theory of elliptic families also shows how arithmetic information may affect chaotic behavior in an elliptical billiard. For instance, *shots from a given point, and having slope of large enough arithmetic height, cannot lead to periodic trajectories* (we tacitly deal here with ellipses and points defined over the algebraic numbers, which implies that the shot-slope is algebraic too if we have periodicity). This kind of implication seems not to have previously appeared in the theory of billiards.

5. Final remarks

Some open issues:

- (1) To prove further cases of the conjectures of Pink and Zilber.
- (2) To achieve effectivity in the counting of rational points appearing in some of the proofs.

This last issue is related to the theory of *o-minimality* in Model Theory. Some crucial recent work towards effectivity is due to Binyamini, and also to Daw, Jones, Schmidt,

- (3) To prove finiteness in families where also the degrees vary.

Some of the methods from *o-minimality* have been developed (by Cluckers, Comte, Forey, and Loeser) in the p -adic context, and already applied by Chambert-Loir and Loeser 2017.

One expects here further applications.

6. References

I have realized that giving references for all the topics that we have touched would lead to a very long list, with some difficult choices and a heavy risk of leaving out something relevant. So, I have decided to quote just two of my own publications on these subjects, whose union contains a relevant quantity of references.

- (1) The book [1] on Unlikely Intersections was written about 10 years ago: much work has appeared later, but the book contains an account of a substantial part of the contents of these notes, and many references.
- (2) The more recent survey paper [2] contains further descriptions and more updated bibliography with respect to the former reference.

Acknowledgements. I thank the Scientific Committee of the 8th ECM for the appreciated invitation to deliver a plenary lecture, at the basis of these notes. I also thank all the Organizers for their excellent job, their hospitality and their precious help in several circumstances.

As to the present notes, I thank Yves André, Pietro Corvaja, David Masser, Barry Mazur, Zeev Rudnick, Peter Sarnak, Jean-Pierre Serre, and Don Zagier for their kind attention and several relevant comments.

References

- [1] U. Zannier, *Some Problems of Unlikely Intersections in Arithmetic and Geometry*. Ann. of Math. Stud. 181, Princeton University Press, Princeton, NJ, 2012 Zbl [1246.14003](#) MR [2918151](#)
- [2] U. Zannier, Some specialization theorems for families of abelian varieties. *Münster J. Math.* **13** (2020), no. 2, 597–619 Zbl [1455.14087](#) MR [4130694](#)

Umberto Zannier

Scuola Normale Superiore, 7 Piazza dei Cavalieri, 56126 Pisa, Italy; umberto.zannier@sns.it