TORSION POINTS ON CURVES

Andrew Granville
Université de Montréal

Zeev Rudnick
Tel-Aviv University

1. Introduction

One of the themes of the summer school is the distribution of “special points” on varieties. In Heath-Brown’s lectures we study rational points on projective hyper-surfaces; in Ullmo’s course we study Galois orbits and Duke’s lectures deal with CM-points on the modular curve. This lecture concerns one of the earliest examples, namely torsion points on group varieties.

DEFINITION 1.1. For a group A, the torsion points are

$$\mathrm{Tor}(A) = \{x \in A : x^n = 1 \text{ for some } n \geq 1\}$$

(we write the group law as multiplication).

If A is abelian then $\mathrm{Tor}(A)$ is a subgroup of A.

EXAMPLES.

(i) The multiplicative group $A = \mathbf{G}_m$ is the algebraic group whose points over a field are the nonzero elements of the field. Then for any field K, $\mathrm{Tor}\,\mathbf{G}_m(K)$ are the roots of unity contained in K.

(ii) $A = \mathbf{G}_m \times \mathbf{G}_m$ then $\mathrm{Tor}(A) = \mathrm{Tor}(\mathbf{G}_m) \times \mathrm{Tor}(\mathbf{G}_m) = \{(x, y) : x, y \in K$ are roots of unity$\}$.

(iii) Let A be an elliptic curve. Over the complex numbers we can uniformize A as $A = \mathbb{C}/L$ where L is a lattice. Then $\mathrm{Tor}(A(\mathbb{C})) = \mathbb{Q} \otimes L/L$.

More generally we can study *division points*:

DEFINITION 1.2. If $\Gamma \subset A$ is a finitely generated group, let

$$\mathrm{Tor}(A, \Gamma) = \{x \in A : x^n \in \Gamma \text{ for some } n \neq 0\}$$

Thus $\mathrm{Tor}(A, \{1\}) = \mathrm{Tor}(A)$ are the torsion points of A.

Motivated by Mordell's conjecture, Lang (Lang, 1965) made the following

CONJECTURE A. *If V is irreducible curve on an abelian group variety (e.g., $A = (\mathbf{G}_m)^n$ or an abelian variety) and $\Gamma \subset A$ is a finitely generated subgroup such that $\mathrm{Tor}(A, \Gamma) \cap V$ is infinite, then V is a translate of a subgroup of A by a division point.*

See Ullmo's lectures (Ullmo, 2006) for the statement of the Manin–Mumford conjecture, which generalizes this statement, and the survey (Tzermias, 2000) for more background.

The first instance of Lang's conjecture is for torsion points on $(\mathbf{G}_m)^r$, which turns out to be quite elementary. We will present two proofs of Lang's conjecture for that case.

2. A Proof Using Galois Theory

The first proof is that which appears in the original paper by Lang (Lang, 1965) where it is attributed to Ihara, Serre and Tate. The result is

THEOREM 2.1. *Let $V/\mathbb{C}$ be an irreducible curve in $A = \mathbf{G}_m \times \mathbf{G}_m$. If V contains infinitely many torsion points then V is a translate of a subgroup of $A = \mathbf{G}_m \times \mathbf{G}_m$ by a torsion point, i.e.,*

$$V = \{(x, y) : x^r = \zeta y^s\}$$

for some root of unity ζ.

To highlight the ideas we will only consider a special case: $V \subset \mathbf{G}_m \times \mathbf{G}_m$ is a rational curve of the forms $\{(f(t), g(t))\}$ where f and g are polynomials, which for added simplicity we assume to have rational coefficients: $f, g \in \mathbb{Q}[t]$. Then

$$V \cap \mathrm{Tor}(A) = \{(f(t), g(t)) \text{ are both roots of unity}\}$$

The subgroups of $\mathbf{G}_m \times \mathbf{G}_m$ are $\{(x, y) : x^r = y^s\}$ for some integers r, s. So we need to show

THEOREM 2.2. *Let $f, g \in \mathbb{Q}[t]$ be polynomials. If there are infinitely many values of t for which both $f(t)$ and $g(t)$ are roots of unity then there are nonzero integers $r, s \neq 0$ so that $f^r = g^s$.*

Proof. We assume there are infinitely many t so that both $f(t)$, $g(t)$ are roots of unity and want to force the relation $f^r = g^s$.

Take $n \gg 1$ so that there is some z_1 with

$$f(z_1) = \zeta_n^\alpha, \quad g(z_1) = \zeta_n^\beta$$

where ζ_n denotes a primitive nth root of unity and that this is the minimal way of writing such an expression, that is $\gcd(n, \alpha, \beta) = 1$ (exercise). Note that $z_1 \in \overline{\mathbb{Q}}$ is algebraic. Then we have a relation

$$f(z_1)^\beta = g(z_1)^\alpha$$

(and both sides equal $\zeta_n^{\alpha\beta}$), but this relation holds for only *one* point z_1 and we want it to hold for *all* points z.

Now apply the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, which acts *transitively* on the primitive n-th roots of unity (see Ullmo's lectures (Ullmo, 2006)). Hence if σ_j is a Galois automorphism so that $\sigma_j(\zeta_n) = \zeta_n^j$, $\gcd(j, n) = 1$ and $z_j := \sigma_j(z_1)$ then because we assume f, g have rational coefficients we get

$$\sigma_j(f(z_1)) = f(\sigma_j(z_1)) = f(z_j), \quad \sigma_j(g(z_1)) = g(z_j)$$

and so

$$f(z_j)^\beta = \sigma_j(f(z_1)) = \zeta_n^{j\alpha\beta} = g(z_j)^\alpha.$$

Now we have the relation $f^\beta = g^\alpha$ holding for $\phi(n)$ distinct points[1] rather than just one point (exercise: why are the points z_j distinct?). However we still need it to hold for *all* z.

Consider the polynomial

$$F(t) = f(t)^\beta - g(t)^\alpha.$$

It has $\phi(n)$ distinct roots so if $\deg F < \phi(n)$ then we would have $F \equiv 0$ as required. Now if $F \neq 0$ then

$$\deg F = \max(\beta \deg f, \alpha \deg g)$$

can be as large as const$\cdot n$. which is still (slightly) too big relative to $\phi(n)$.

The remedy is to raise the relation $f(z_j)^\beta = g(z_j)^\alpha = \zeta_n^{\alpha\beta}$ to an m-th power:

$$f(z_j)^{m\beta} = g(z_j)^{m\alpha}$$

(both sides equal $\zeta_n^{m\alpha\beta}$). We get a new polynomial $f^{m\beta} - g^{m\alpha}$ with $\phi(n)$ distinct roots; it looks like we raised the degree which is certainly useless! However, since $f(z_j)$, $g(z_j)$ are nth roots of unity, we have $f(z_j)^n = 1 = g(z_j)^n$ and if we substitute

$$m\beta \equiv r \bmod n, \quad m\alpha \equiv s \bmod n$$

with $|r|, |s| \leq n/2$ then we find $f(z_j)^r = g(z_j)^s$ for all j coprime to n. This is still not useful as we have just showed that $\deg F \leq \max(\deg f, \deg g)n/2$ instead of showing that $\deg F < \phi(n)$. However we will be done if we can show that there is some $m \geq 1$ so that the residues $(m\beta, m\alpha) \bmod n$ are both small! This is given by the following

[1] $\phi(n)$ is the number of residues coprime to n

EXERCISE. Given a primitive vector $(\alpha, \beta) \in (\mathbb{Z}/n\mathbb{Z})^2$, that is $\gcd(\alpha, \beta, n) = 1$, there is some $1 \le m \le n$ so that both residues $m\alpha \bmod n$ and $m\beta \bmod n$ are at most $n^{2/3}$ (and are different than $(0, 0) \bmod n$).

See Venkatesh's lecture (Venkatesh, 2006) and (Strombergsson and Venkatesh, 2005) where it is shown that typically the size of both residues is about $\sqrt{n}$.

Consequently we find a relation $f(z_j)^r = g(z_j)^s$ with $|r|, |s| < n^{2/3}$ and hence $\deg F \ll n^{2/3}$. Since $\phi(n) \gg n^{1-\epsilon}$ for all $\epsilon > 0$, the assumption that there are infinitely many torsion points (that is we can take n arbitrarily large) implies the identity $f^r = g^s$ as required.

3. Polynomials Vanishing at Roots of Unity

In this section we present a proof of the following strong version of Lang's conjecture for torsion points on a variety V in $\mathbb{C}^m$. We denote by $\mathbb{U}_{\text{tors}}$ be the set of roots of unity.

COROLLARY 3.1. *Let V be an algebraic variety embedded in $\mathbb{C}^m$. There exists an explicitly computable, finite list $\mathcal{B}$ of ℓ_B-by-m integer matrices B, with each $\ell_B \ge 1$, such that if $\zeta \in V(\mathbb{U}_{\text{tors}})$ then $\zeta \in \bigcup_{B \in \mathcal{B}} W_B(\mathbb{U}_{\text{tors}})$ where $W_B = \bigcap_{j=1}^{\ell_B} \{\zeta : \zeta_1^{b_{j,1}} \zeta_2^{b_{j,2}} \cdots \zeta_m^{b_{j,m}} = 1\}$.*

It is not difficult to give an explicit description of $W(\mathbb{U}_{\text{tors}})$—see at the end.

To prove this result we shall develop a simple understanding of vanishing sums of roots of unity– see (Conway and Jones, 1976) and (Lenstra, 1979) for far more. We begin by considering a linear form $a_1X_1 + a_2X_2 + \cdots + a_kX_k$ where each a_i is an integer. We are interested in finding all sets $(\xi_1, \xi_2, \ldots, \xi_k) \in \mathbb{U}_{\text{tors}}^k$ such that $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$. We call such a sum *minimal* if no proper vanishing sums of roots of unity subsum equals zero (that is, there does not exist a proper subset I of $\{1, \ldots, k\}$ for which $\sum_{i \in I} a_i\xi_i = 0$); it occurs no loss of generality in our calculations to partition any such sum into minimal subsums. Given any such minimal solution there are *equivalent* solutions $(\xi\xi_1, \xi\xi_2, \ldots, \xi\xi_k)$ for any root of unity ξ. Two solutions are *equivalent* if they can be partitioned (in the same way) into minimal subsums, where the the corresponding subsums are equivalent.

For any set $(\xi_1, \xi_2, \ldots, \xi_k) \in \mathbb{U}_{\text{tors}}^k$ there is a minimal $n = n(\xi_1, \xi_2, \ldots, \xi_k)$ for which $(\xi_i/\xi_j)^n = 1$ for each pair $1 \le i, j \le k$. Note that any minimal sum $\sum_{i=1}^k a_i\xi_i = 0$ is thus equivalent to a minimal solution $\sum_{i=1}^k a_i\xi_i' = 0$ where each $(\xi')^n = 1$, with $n = n(\xi_1, \xi_2, \ldots, \xi_k)$. Our key result is the following:

PROPOSITION 3.2. *Suppose that $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$ is minimal. Then $n(\xi_1, \xi_2, \ldots, \xi_k)$ is squarefree, and if prime p divides n then $p \le k$. Therefore n divides $N_k := \prod_{p \le k} p$.*

Given non-zero integers $a_1, a_2, \ldots, a_k$, let $\mathbb{X} = \mathbb{X}(a_1, \ldots, a_k)$ be the set

$$\{(\xi_1, \ldots, \xi_k) : \xi_j^{N_k} = 1 \text{ for each } j, \text{ and } a_1\xi_1 + \ldots + a_k\xi_k = 0\},$$

which is finite and computable, simply by trying all possible values for each ξ_j. One consequence of Proposition 3.2 is the following result:

COROLLARY 3.3. *Suppose $a_1, \ldots, a_k \in \mathbb{Z}^*$. For given $(\xi_1, \xi_2, \ldots, \xi_k) \in \mathbb{U}_{\text{tors}}^k$ we have $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$ if and only if $(\xi_1, \xi_2, \ldots, \xi_k)$ is equivalent to an element of $\mathbb{X}$.*

Proof. Given $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$, split the sum up into minimal subsums, each one of which (according to the remarks above) is equivalent to one where each ξ_i is an nth root of unity. Moreover n divides $N_\ell = \prod_{p \leq \ell} p$ by Proposition 3.2, where ℓ is the length of the subsum, and the result follows since $\ell \leq k$. On the other hand if $(\xi_1, \xi_2, \ldots, \xi_k)$ is equivalent to an element of $\mathbb{X}$ then $a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k = 0$ by the definition of $\mathbb{X}$.

With that preparation we can prove Corollary 3.1:

Proof of Corollary 3.1. An algebraic variety can be described as the set of points in $\mathbb{C}^m$ satisfying certain equations with algebraic coefficients; and this is a subset of the algebraic variety given by the set of points in $\mathbb{C}^m$ satisfying the norms of these equations, which are equations with integer coefficients. So without loss of generality we will assume the coefficients of the polynomials defining V are integers.

Now suppose that

$$f_j(x_1, \ldots, x_m) = \sum_{i=1}^{k_j} a_{j,i} x_1^{s_{j,i,1}} x_2^{s_{j,i,2}} \cdots x_m^{s_{j,i,m}} \in \mathbb{Z}[x_1, \ldots, x_m]$$

for $1 \leq j \leq J$. We are interested in $\zeta \in \mathbb{U}_{\text{tors}}^m$ for which $f_j(\zeta) = 0$ for each j; evidently these induce solutions to

$$a_{j,1}\xi_{j,1} + a_{j,2}\xi_{j,2} + \cdots + a_{j,k_j}\xi_{j,k_j} = 0$$

with each $\xi_{j,i} = \zeta_1^{s_{j,i,1}} \zeta_2^{s_{j,i,2}} \cdots \zeta_m^{s_{j,i,m}}$. Now each of these vanishing sums can be partitioned into minimal vanishing subsums; let us relabel one of these minimal vanishing subsums to be $a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = 0$. As we saw in Proposition 3.2, each $\xi_r/\xi_1 = \zeta_1^{s_{r,1}-s_{1,1}} \zeta_2^{s_{r,2}-s_{1,2}} \cdots \zeta_m^{s_{r,m}-s_{1,m}}$ must be an N_kth root unity, so $\zeta_1^{b_{r,1}} \zeta_2^{b_{r,2}} \cdots \zeta_m^{b_{r,m}} = 1$ where $b_{r,j} = N(s_{r,j} - s_{1,j})$ for each j. We get sets of such vectors b_r for each minimal vanishing subsum (and from each f_j) and we can concatenate these all together to form one large matrix B (with, say, ℓ rows), and so $\zeta \in W_B(\mathbb{U}_{\text{tors}})$.

Finally, since there are only finitely many possible partitions into minimal subsums, the set $\mathcal{B}$ of such matrices B, is finite and computable.

Proof of Proposition 3.2. Write each $\xi_j = e(k_j/n)$ with $0 \le k_j \le n-1$.

Suppose that integer r divides n, and let $\beta_j \equiv k_j \pmod{n/r}$ with $0 \le \beta_j \le n/r - 1$, and $\gamma_j = (k_j - \beta_j)/(n/r)$ so that $0 \le \gamma_j \le r-1$. Thus $\xi_j = e(\beta_j/n)e(\gamma_j/r)$. Now, for each $0 \le i \le r-1$ and $0 \le \ell \le n/r - 1$, let $A_{i,\ell}$ be the sum of the a_j with $\beta_j = \ell$ and $\gamma_j = i$ so that

$$0 = a_1\xi_1 + a_2\xi_2 + \ldots + a_k\xi_k = \sum_{j=0}^{k} a_j e(\beta_j/n)e(\gamma_j/r) = \sum_{\ell=0}^{n/r-1} \left(\sum_{i=0}^{r-1} A_{i,\ell} e(i/r) \right) e(\ell/n).$$

Let $r = r(n) = \prod_{p|n} p$ and recall that $[\mathbb{Q}(e(1/n)) : \mathbb{Q}(e(1/r))] = n/r$ (by elementary Galois theory) and so $e(\ell/n)$, $0 \le \ell \le n/r - 1$ are linearly independent over $\mathbb{Q}(e(1/r))$. In particular this implies that each of the subsums $\sum_{i=0}^{r-1} A_{i,\ell}\, e(i/r) = 0$ above, which contradicts our assumption of minimality, unless $A_{i,\ell} = 0$ for all i for all $\ell \ne \ell_0$ for some ℓ_0; in other words $\beta_j = \ell_0$ for all j. But then $\xi_i/\xi_j = e(\ell_0/n)e(\gamma_j/r)/e(\ell_0/n)e(\gamma_j/r) = e((\gamma_i - \gamma_j)/r)$ and so $n(\xi_1, \xi_2, \ldots, \xi_k)$ divides r. Thus $n = r(n)$ is squarefree.

Since n is squarefree we may write $n = mp$ with $(m, p) = 1$. Then, by the Chinese Remainder theorem there exists $0 \le \beta_j \le p-1$ and $0 \le \gamma_j \le m-1$ such that $k_j \equiv m\beta_j \pmod{p}$ and $k_j \equiv p\gamma_j \pmod{m}$ and thus $\xi_j = e(\beta_j/p)e(\gamma_j/m)$. Letting $A_{i,\ell}$ now be the sum of the a_j with $\beta_j = \ell$ and $\gamma_j = i$ we obtain

$$0 = a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k = \sum_{j=0}^{k} a_j e(\beta_j/p)e(\gamma_j/m) = \sum_{\ell=0}^{p-1} \left(\sum_{i=0}^{m-1} A_{i,\ell} e(i/m) \right) e(\ell/p).$$

Recall that $[\mathbb{Q}(e(1/n)) : \mathbb{Q}(e(1/m))] = p - 1$ (by elementary Galois theory), so that the only linear dependencies between $e(\ell/p)$, $0 \le \ell \le p-1$, over $\mathbb{Q}(e(1/m))$, are multiples of $\sum_{\ell=0}^{p-1} e(\ell/p) = 0$. Therefore from the equation above we see that $\sum_{i=0}^{m-1} A_{i,\ell} e(i/m) = \lambda$ for some $\lambda \in \mathbb{Q}(e(1/m))$. Evidently $\lambda \ne 0$ else, by the argument from the paragraph above we see that $n \mid m$. Therefore for each ℓ there exists i with $A_{i,\ell} \ne 0$ and in particular some $j = j_\ell$ with $\beta_{j_\ell} = \ell$; and so $p \le k$ as claimed.

3.1. DETERMINING $W_B(\mathbb{U}_{\text{tors}})$

Suppose that the ℓ-by-m integer matrix B is given and we write each $\zeta_j = e(v_j)$, so the points in W_B correspond exactly to those $v \in (\mathbb{Q}/\mathbb{Z})^m$ satisfying $Bv \equiv 0 \pmod{1}$. Note that if $y \in B^{\perp}(\mathbb{Q}) \pmod{1}$ then $By \equiv 0 \pmod{1}$, so we call two solutions v, v' *equivalent* if $v - v' \in B^{\perp}(\mathbb{Q}) \pmod 1$. We will prove that there are no more than finitely many inequivalent solutions, which are effectively computable:

We wish to use the tools of linear algebra to solve this equation but there are many zero divisors in $\mathbb{Q} \pmod 1$ (indeed if $a/q \in \mathbb{Q}$ then $q \cdot (a/q) \equiv 0 \pmod 1$), so we avoid any division! In Gaussian elimination one diagonalizes as much of the matrix as possible, dividing non-zero elements in a given row by the "pivot element" (that is if $B_{1,1} \neq 0$ is the pivot element then one replaces the current row i by the current row i minus $B_{i,1}/B_{1,1}$ times the first row). This can be reworked to avoid division simply by introducing multiples (that is we replace the current row i by $B_{1,1}$ times the current row i minus $B_{i,1}$ times the first row). Note that any solution of the original linear algebra problem is also a solution of the new problem; and vice-versa whenever $B_{1,1}$ is invertible, though if this is not so (as may be the case here) this process may well introduce several bogus solutions. Nonetheless at the end of the Gaussian elimination process we have an l-by-m integer matrix B' (with $l \leq \ell$ after deleting rows of 0s), in which the left-most l-by-l submatrix is diagonal with non-zero diagonal entries (if necessary by swapping various rows and columns), for which $B'v \equiv 0 \pmod 1$. Solving this is easy: there are $m - l$ free variables $v_{l+1}, v_{l+2}, \ldots, v_m$ and, writing $\beta_i = B'_{i,i}$, we have $v_i \equiv (u_i - \sum_{j=l+1}^{m} B'_{i,j} v_j)/\beta_i \pmod 1$, where u_i is any integer with $0 \leq u_i \leq \beta_i - 1$.

For $l + 1 \leq j \leq m$ let y_j be the vector with ith entry $-B'_{i,j}/\beta_i$ for $1 \leq i \leq \ell$, and $\delta_{i,j}$ otherwise (where δ is the Dirac delta function). The solutions to $B'v \equiv 0 \pmod 1$ all take the form $v = u + \sum_{j=l+1}^{m} v_j y_j$ where $u \in U'$ a finite computable set. If we trace through the proof above then we find that $By_j = 0$ for each j, that is each $y_j \in B^{\perp}$. Thus there is a set U of representatives of the equivalence classes of solutions inside U' which can be determined by testing whether they satisfy $Bu \equiv 0 \pmod 1$.

References

Conway, J. H. and Jones, A. J. (1976) Trigonometric diophantine equations on vanishing sums of roots of unity, *Acta Arith.* **30**, 229–240.

Lang, S. (1965) Division points on curves, *Ann. Mat. Pura Appl. (4)* **70**, 229–234.

Lenstra, Jr., H. W. (1979) Vanishing sums of roots of unity, In *Proc. Bicentennial Congress Wiskundig Genootschap*, Vol. 101 of *Math. Centre Tracts*, Vrije Univ. Amsterdam, 1978, pp. 249–268, Math. Centrum, Amsterdam.

Strombergsson, A. and Venkatesh, A. (2005) Small solutions to linear congruences and Hecke equidistribution, *Acta Arith.* **118**, 41–78.

Tzermias, P. (2000) The Manin–Mumford conjecture: a brief survey, *Bull. London Math. Soc.* **32**, 641–652.

Ullmo, E. (2006) Manin–Mumford, André–Oort, the equidistribution point of view, in this book.

Venkatesh, A. (2006) Spectral theory of automorphic forms: a very brief introduction, in this book.

INDEX