Rend. Sem. Mat. Univ. Pol. Torino
Vol. 53, 3 (1995)
Number Theory

L. Caporaso

COUNTING RATIONAL POINTS ON ALGEBRAIC CURVES

Abstract. We describe recent developments on the problem of finding examples of algebraic curves of genus at least 2 having the largest possible number of rational points. This question is related to the Conjectures of Lang on the distribution of rational points on the varieties of general type.

1. The question

To formulate our main question let us fix the field K and consider curves (that is, smooth, irreducible, projective, algebraic curves) having given genus $g \geq 2$, then our challenge is

How many K-rational points can a curve of genus g defined over K have?

This very naive question was asked by J. Harris, B. Mazur and myself for the first time during the Summer of 1993, for the reasons explained in the following section. At that time, as we found out, there were very few examples of curves with many rational points. The record holding curves in low genus over the rational numbers had been discovered by Brumer, they had 144 points in genus 2 and 72 points in genus 3. This paper will report on the progress that has been made on this and on other related subjects.

2. Motivation: Lang conjectures and Uniformity conjectures

Special interest in the question described above is a consequence of results obtained in [2]. One goal of such a paper is to prove that the Lang Diophantine conjectures are directly related to the Uniformity conjectures.

The Lang conjectures try to generalize Faltings' Theorem (that a curve of genus at least 2 has finitely many rational points) to varieties of higher dimension. Here are the statements of two of them (maybe the best known among all).

WEAK LANG CONJECTURE. *Let X be a variety of general type defined over K. Then the set of K-rational points of X is not dense in X with respect to the Zariski topology.*

Of course for the above statement not to be obviously false, we must assume that X has positive dimension. Also we use the terminology "variety of general type" for irreducible and reduced varieties only. Recall that a smooth variety X is of general type if some power of its canonical bundle has enough sections to effect an embedding of some open subset of X in projective space. A general variety is said to be of general type if some desingularization of it is of general type. For example, a smooth curve is of general type if and only if it has genus greater than 1. Therefore the Weak Lang conjecture is true in dimension 1.

STRONG LANG CONJECTURE. *Let X be a variety of general type defined over K. Then there exists a proper closed (in the Zariski topology) subvariety Ξ of X such that for any number field L containing K all of the L-rational points of X are contained in Ξ with the exception of finitely many of them.*

The only reason why we mentioned these two conjectures in this note, is that they imply two "Uniformity" conjectures about the distribution of rational points on curves. This is proved in [2], Theorems 1.1 and 1.2, the statements of which we recall:

UNIFORM BOUND THEOREM. *The Weak Lang Conjecture implies that, for every number field K and for every $g \geq 2$, there exists a number $\mathrm{B}(K, g)$ such that no curve of genus g defined over K has more than $\mathrm{B}(K, g)$ points defined over K.*

We remark that the result of this theorem has been improved by D. Abramovich and P. Pacelli (see [1] and [7]) who showed that if such a bound $\mathrm{B}(K, g)$ exists, then it only depends on the degree of the field extension (over the rational numbers, for example, or over any fixed base field), rather than the field itself. The second implication between Lang conjectures and Uniformity conjectures is stated in the following theorem:

UNIVERSAL GENERIC BOUND THEOREM. *The Strong Lang Conjecture implies that, for every $g \geq 2$, there exists a number $\mathrm{N}(g)$ such that for any number field K there are only finitely many K-isomorphism classes of curves of genus g defined over K having more than $\mathrm{N}(g)$ K-rational points.*

In [3] we therefore asked: How many rational points can an algebraic curve have? Having those two results in mind, the goal is to obtain lower bounds on $\mathrm{B}(K, g)$ and $\mathrm{N}(g)$. In [3] we described a number of methods to approach such questions; some of these methods (and in fact the most efficient of them) are due to A. Brumer, N. Elkies and J.F. Mestre. This report can be viewed as an update of [3]; it will contain, among other things, a list of new and old but standing records, together with a description of a beautiful technique due to A. Brumer, that, to my knowledge, has not been described in any paper. The spirit with which this paper (as well as [3]) is thought, is to present the various approaches that have been used, to stimulate more research on the subject. It is fair to say that the results

obtained so far fail to shed much light on the main questions: how does the number of rational points that a curve has vary, as we vary the curve? Is it bounded once we fix the genus and the number field? What if we fix the genus only?

We will soon give a better picture of what the state of the knowledge is, regarding to these questions. Before doing that, I will explain the reason why the Lang conjectures imply the Uniformity conjectures. This will be the content of the coming section; those who are only interested in the arithmetic aspect can just skip it.

3. The Correlation Theorem and its implications in arithmetic geometry

The two theorems stated before, that is, the fact that the Lang conjectures imply certain Uniformity conjectures, are really corollaries of what we call the "Correlation" Theorem (Theorem 1.3 in [2]), the proof of which occupies most of [2]. This is a purely algebro-geometric result, but we gave it such a name because we had in mind to use it together with the Lang conjectures. A more detailed explanation of this point of view will be given after its statement.

THEOREM (Correlation). *Let $X \to B$ be a proper morphism of reduced and irreducible varieties whose general fiber is a curve of genus at least 2. Then for n big enough, the n-th fiber product of X over B admits a dominant rational map h to a positive dimensional variety of general type W. If X is defined over the field K, then W and h are also defined over K.*

Now let us see what is the idea to show, for example, that the Correlation Theorem together with the Weak Lang conjecture imply the existence of a (finite) bound on the number of rational points that any curve of fixed genus can have. Let us look at a family $X \to B$ of curves of genus 2 or more (we might very well view this as a family containing all isomorphism classes of such curves). Faltings Theorem tells us that each fiber has finitely many rational points; the "problem" is that, as we vary the fiber, these finite sets seem to be completely unrelated and we do not have any control over their cardinality, which we consider as a function on the base B. But, if the Lang conjecture holds, we can apply it to the variety W of the Correlation Theorem, whose rational points will then all be contained in a proper closed subvariety Z. This will give us in turn that the rational points of the n-th order fiber product of X over B are also all contained in a proper closed subvariety Z'. Finally, the algebraic equations defining Z' can then be viewed as functions "co-relating" n-uples of rational points of the curves of our family. In this way we get a hold of how the sets of rational points vary when varying the fiber, and we can show that their cardinality is a bounded function.

4. Counting rational points: an update on current records.

It is clear what the strategy should be to find lower bounds on the numbers $\mathrm{B}(K,g)$ (the uniform bound) and $\mathrm{N}(g)$ (the universal bound). For the first, one simply has to construct examples of curves over a certain field, in a way that there are many visible rational points. For the second, one has to produce families of non-isomorphic curves, (that is, infinitely many curves that are not isomorphic over the algebraic closure of $\mathbb{Q}$) of given genus, defined over some number field, with as many rational points as possible.

First of all we notice that, as of today, nobody was able to disprove any of the above conjectures, that is, nobody could prove that $\mathrm{B}(K,g)$ and $\mathrm{N}(g)$ are infinite. On the other hand, almost all the results of [3] have been significantly improved. Now the readers could find this either good or bad, depending on their faith in the Lang conjectures! In any case, here is the state of the art.

Bounds on $\mathrm{B}(K,g)$

$\mathrm{B}(\mathbb{Q},2) \geq 588$ (L. Kulesz)

improving the 144 appearing in [3]. This record holding curve is given by the equation

$$y^2 = 278271081x^2(x^2-9)^2 - 229833600(x^2-1)^2.$$

It has at least 12 automorphisms (notice that $588 = 12 \cdot 49$). It is also of interest to look at curves without extra automorphisms, for the case of genus 2 the record is held by C.Stahlke with the curve

$$y^2 = 9703225x^6 - 9394700x^5 + 152200x^4 + 1124745x^3 + 119526x^2 - 42957x + 2061$$

which has at least 306 rational points.

$\mathrm{B}(\mathbb{Q},3) \geq 112$ (W. Keller-L. Kulesz)

which improves a long standing record of 72. The curve is

$$y^2 = 48397950000(x^2+1)^4 - 939127350499(x^3-x)^2,$$

having at least 16 automorphisms.

$\mathrm{B}(\mathbb{Q},4) \geq 126$ (N. Elkies)

$\mathrm{B}(\mathbb{Q},g) \geq 8g+12$ (J.F. Mestre)

The results above are obtained by a method of J.F. Mestre, for a partial description of which the reader can read Section 5.3 of [3]. Also, people used a C-program due to N.Elkies to quickly find rational points on curves.

Bounds on $\mathrm{N}(g)$

With the exception of the genus 2 case, the bounds in the list below are obtained with families defined over cyclotomic extensions of the rationals. For a completely geometric proof of the fact that $\mathrm{N}(2) \geq 128$, see Section 4 of [3].

g	2	3	4	5	9	10	45
N(g)	≥ 128	≥ 100	≥ 126	≥ 146	≥ 180	≥ 192	≥ 781

Except for the first one all these bounds are due to N. Elkies. We also have in general

$$\mathrm{N}(g) \geq 16(g+1) \qquad \text{(A. Brumer)}$$

This last result will be proved in the next section. Before doing that, I want to express my gratitude to Noam Elkies, for his valuable help in compiling the list of records above.

5. The method of A. Brumer

The result that one obtains using this technique is

PROPOSITION (Brumer). *Let ζ be a primitive $(g+1)$-th root of* 1, *then*

(i) $\mathrm{B}(\mathbb{Q}(\zeta), g) \geq 16(g+1)$

(ii) $\mathrm{N}(g) \geq 16(g+1)$.

Proof. These bounds are obtained with hyperelliptic curves having big automorphism group. Consider the following family of hyperelliptic curves:

$$C_{a,b} = V\left(y^2 - a(x^n+1)^2 - bx^n\right)$$

where a and b are nonzero numbers varying in the chosen base field. (The above notation means that $C_{a,b}$ is given as the locus of points in the plane satisfying the equation $y^2 = a(x^n+1)^2 - bx^n$.)

The curves $C_{a,b}$ are therefore hyperelliptic of genus $g = n-1$. The first observation is that $C_{a,b}$ is given as a double cover of the projective line $\mathbb{P}^1$, with ramification points lying on two concentric rings. Therefore, the curves have two natural types of automorphisms, besides the hyperelliptic involution. Namely, we have the involution that exchanges the two rings, which is given by

$$(x,y) \to \left(\frac{1}{x}, \frac{y}{x^n}\right);$$

and we have the rotations that leave the ramification locus invariant; there are n of them, given by

$$(x,y) \to (\zeta^i x, y)$$

where ζ is a primitive n-th root of 1. We can therefore conclude that

$$|\mathrm{Aut}(C_{a,b})| \geq 4(g+1)$$

and these automorphisms are all defined over the field $\mathbb{Q}(\zeta)$.

After this preliminary analysis, we start looking for rational points. If we find a number m of them (defined over the rationals, say), by looking at their orbits under the automorphism group, we should get $4m(g+1)$ points defined over $\mathbb{Q}(\zeta)$. And now, here is the way we approach this problem: for every fixed $m = 1, 2, 3...$ we want to find a and b in $\mathbb{Q}$ such that $C_{a,b}$ contains m points defined over $\mathbb{Q}$.

Now, for each m, our question translates into a linear problem having a and b as unknowns, all we will have to do is to find nontrivial solutions for m as large as possible. In fact, if (x_i, y_i) is a rational point of $C_{a,b}$ for $i = 1, ..., m$, we get a system of m linear equations in a and b:

$$y_i^2 = a(x_i^n + 1)^2 + bx_i^n \quad i = 1, ..., m.$$

Let A be the $m \times 3$ matrix of the coefficients of the system:

$$A = \begin{pmatrix} \vdots & \vdots & \vdots \\ y_i^2 & (x_i^n + 1)^2 & x_i^n \\ \vdots & \vdots & \vdots \end{pmatrix}.$$

then the system has non-trivial solutions if and only if the the rank of A is less than 3.

Hence the cases $m = 1, 2$ are easily seen to have infinitely many solutions. The first interesting case is $m = 3$; actually, we might very well skip to the case $m = 4$, which gives the proof of the Proposition. The only reason to write down the case $m = 3$ is that it is quite simple and it makes a useful warm-up. We have the condition $detA = 0$, which can be viewed as the equation of a conic in the projective plane $\mathbb{P}^2_F$ over the field $F = K(x_1, x_2, x_3)$. It is easy to find a $\mathbb{Q}$-rational point on such a conic: just take the point $y_i = x_i^n + 1$ with $i = 1, 2, 3$. Such an obvious rational point corresponds of course to the uninteresting solution $b = 0$, but its existence guarantees that the conic has infinitely many other rational points. This shows that for a generic choice of x_1, x_2, x_3 in $\mathbb{Q}$ there exist infinitely many pairs of rational numbers (a, b) such that the curve $C_{a,b}$ contains 3 rational points. In fact one can see that there are infinitely many isomorphism classes of such curves, so that a first result is that

$$\mathrm{N}(g) \geq 12(g+1)$$

but one can do better, let us in fact treat the case $m = 4$ in a similar fashion. Now, to impose that the rank of the matrix A be equal to 2, gives two equations in y_1, y_2, y_3, y_4 that we interpret as two quadrics in the projective space $\mathbb{P}^3_F$, where F is the field $F = K(x_1, x_2, x_3, x_4)$. In other words, the locus of points (y_1, y_2, y_3, y_4) for which $rkA \leq 2$ is a curve E obtained as the intersection of two quadrics in projective space; this is an elliptic curve (provided that the pair of quadrics is general), and this particular curve E has some obvious rational points (in fact 8 of them), namely $y_i = \pm(x_i^n + 1)$ with

$i = 1, 2, 3, 4$. Just as in the case $m = 3$, the fact that there are some obvious rational points on E (corresponding to degenerate curves $C_{a,b}$) implies the existence of other rational points, which are obtained by adding the obvious ones on E. We can then conclude that for a generic choice of four rational numbers x_1, x_2, x_3, x_4 there exist rational numbers a and b such that the curve $C_{a,b}$ contains at least four rational points (x_i, y_i) and hence at least $16(g+1)$ points defined over the field $\mathbb{Q}(\zeta)$. This proves the first statement in the Proposition. To finish the proof of (ii) we have to show that this method yields infinitely many curves that are not isomorphic over the algebraic closure of $\mathbb{Q}$. For that, notice that this construction shows that given 4 generic points x_1, x_2, x_3 and x_4 in $\mathbb{P}^1_{\mathbb{Q}}$ we can find a hyperelliptic curve of type $C_{a,b}$ having 4 rational points lying over the x_i. Therefore as $C_{a,b}$ varies we find infinitely many values of the cross-ratio of the set x_1, x_2, x_3, x_4. But now, if there were only finitely many $\overline{\mathbb{Q}}$-isomorphism classes of these curves, we would only find a finite number of values for the cross-ratio of their rational points, and this is impossible, as we have just observed.

REFERENCES

[1] ABRAMOVICH D., *Uniformité des points rationnels des courbes algébriques sur les extensions quadratiques et cubiques.* C.R Acad.Sci.Paris, t.321, S.I, (1995) p. 755-758.

[2] CAPORASO L., HARRIS J., MAZUR B., *Uniformity of rational points.* To appear on Journal of the american mathematical society.

[3] CAPORASO L., HARRIS J., MAZUR B., *How many rational points can a curve have?* The Moduli space of curves, Progress in mathematics n. 129, Birkhäuser (1995) p.13-32.

[4] FALTINGS G., *The general case of Lang's conjecture.* Preprint.

[5] LANG S., *Hyperbolic and diophantine analysis.* Bull.Amer.Math.Soc. **14**, No.2 (1986) p.159-205.

[6] MUMFORD D., *Stability of projective varieties.* L'enseignement Mathématique **23** (1977) p.39-110.

[7] PACELLI P., *Uniform boundedness for rational points.* Ph.D. Thesis, Boston University, (1996).

Lucia CAPORASO
Dipartimento di Matematica
Università di Roma Tor Vergata
00100 Roma, Italy.