ACTA ARITHMETICA
113.1 (2004)

Torsion points on curves and common divisors of $a^k - 1$ and $b^k - 1$

by

NIR AILON (Tel Aviv and Princeton) and ZÉEV RUDNICK (Tel Aviv)

1. Introduction. Let $a, b \neq \pm 1$ be nonzero integers. One of our goals in this paper is to study the common divisors of $a^k - 1$ and $b^k - 1$, specifically to understand small values of $\gcd(a^k - 1, b^k - 1)$. If $a = c^u$ and $b = c^v$ for some integer c then clearly $c^k - 1$ divides $\gcd(a^k - 1, b^k - 1)$ and so for the purpose of understanding small values, we will assume that a and b are *multiplicatively independent*, that is, $a^r \neq b^s$ for $r, s \geq 1$. Further, since $\gcd(a-1, b-1)$ always divides $\gcd(a^k - 1, b^k - 1)$, we will assume that $a - 1$ and $b - 1$ are coprime.

Based on numerical experiments and other considerations, we conjecture:

CONJECTURE A. *If a, b are multiplicatively independent non-zero integers with* $\gcd(a-1, b-1) = 1$, *then there are infinitely many integers $k \geq 1$ such that*

$$\gcd(a^k - 1, b^k - 1) = 1.$$

Note that the condition of multiplicative independence of a and b is not necessary, as the (trivial) example $b = -a$ shows (the gcd is 1 for odd k, and $a^k - 1$ for even k).

A recent result of Bugeaud, Corvaja and Zannier [BCZ] rules out *large* values of $\gcd(a^k - 1, b^k - 1)$. They show that if $a, b > 1$ are multiplicatively independent positive integers then for all $\varepsilon > 0$,

$$\gcd(a^k - 1, b^k - 1) \ll_\varepsilon e^{\varepsilon k} . \tag{1}$$

Their argument uses Diophantine approximation techniques and in particular Schmidt's Subspace Theorem. They also indicate that there are arbitrarily large values of k for which the upper bound (1) cannot be significantly improved.

In the function field case, when we replace integers by polynomials, we are able to prove a strong version of Conjecture A.

2000 *Mathematics Subject Classification*: Primary 11D61.

Theorem 1. *Let $f, g \in \mathbb{C}[t]$ be nonconstant polynomials. If f and g are multiplicatively independent, then there exists a polynomial h such that*

$$\gcd(f^k - 1, g^k - 1) \mid h \tag{2}$$

for any $k \geq 1$. If, in addition, $\gcd(f - 1, g - 1) = 1$, then there is a finite union of proper arithmetic progressions $\bigcup d_i\mathbb{N}$, $d_i \geq 2$, such that for k outside these progressions,

$$\gcd(f^k - 1, g^k - 1) = 1.$$

Note that (2) is a strong form of (1). We derive Theorem 1 from a result proposed by Lang [L1] on the finiteness of torsion points on curves—see Section 2.

We next consider a generalization to the case of matrices. For an $r \times r$ integer matrix $A \in \mathrm{Mat}_r(\mathbb{Z})$, $A \neq I$ (I being the identity matrix), we define $\gcd(A - I)$ as the greatest common divisor of the entries of $A - I$. Equivalently, $\gcd(A - I)$ is the greatest integer $N \geq 1$ such that $A \equiv I \bmod N$. We say that A is *primitive* if $\gcd(A - I) = 1$. Note that $\gcd(A - I)$ divides $\gcd(A^k - I)$ for all k. A similar definition applies to the function field case $A \in \mathrm{Mat}_r(\mathbb{C}[t])$. We will study the behaviour of $\gcd(A^k - I)$ as k varies for a fixed matrix A with coefficients in $\mathbb{Z}$ or in $\mathbb{C}[t]$. If $\det A = 0$ then trivially $\gcd(A^k - I) = 1$ for all $k \geq 1$. So we will henceforth assume that A is nonsingular.

For the case of 2×2 matrices, we will show in Section 3 that if $A \in \mathrm{SL}_2(\mathbb{Z})$ is unimodular and hyperbolic, then $\gcd(A^k - I)$ grows exponentially as $k \to \infty$. However, numerical experiments show that for other matrices, $\gcd(A^k - I)$ displays completely different behaviour. We formulate the following conjecture:

Conjecture B. *Suppose $r \geq 2$ and $A \in \mathrm{Mat}_r(\mathbb{Z})$ is nonsingular and primitive. Also assume that there is a pair of eigenvalues of A that are multiplicatively independent. Then A^k is primitive infinitely often.*

Note that Conjecture B subsumes Conjecture A. It would be interesting to prove an analogue of the upper bound (1) in this setting.

In Section 4 we give an example where we can prove Conjecture B. To describe it, recall that one may obtain integer matrices by taking an algebraic integer u in a number field K and letting it act by multiplication on the ring of integers $\mathcal{O}_K$ of K. This is a linear map and a choice of integer basis of $\mathcal{O}_K$ gives us an integer matrix $A = A(u)$ whose determinant equals the norm of u. We employ this method for the cyclotomic field $\mathbb{Q}(\zeta_p)$ where $p > 3$ is prime and ζ_p is a primitive pth root of unity, and u is a nonreal unit. We show:

THEOREM 2. *Let u be a nonreal unit in the extension $\mathbb{Q}(\zeta_p)$, and $A(u) \in \mathrm{SL}_{p-1}(\mathbb{Z})$ be the corresponding matrix. Then $A(u)^k$ is primitive for all $k \not\equiv 0 \bmod p$.*

In the function field case, we have a strong form of Conjecture B, which generalizes Theorem 1:

THEOREM 3. *Let A be a nonsingular matrix in $\mathrm{Mat}_r(\mathbb{C}[t])$. Assume that either*

(1) *A is not diagonalizable over the algebraic closure of $\mathbb{C}(t)$, or*
(2) *A has two eigenvalues that are multiplicatively independent.*

Then there exists a polynomial h such that $\gcd(A^k - I) \mid h$ for any k. If, in addition, A is primitive, then A^k is primitive for all k outside a finite union of proper arithmetic progressions.

Acknowledgements. We would like to thank Umberto Zannier for useful discussions and the referee for suggesting several improvements. Some of the results were part of the first named author's M.Sc. thesis [Ai] at Tel Aviv University. The work was partially supported by the Israel Science Foundation, founded by the Israel Academy of Sciences and Humanities.

2. Proof of Theorem 1. To prove the theorem, we will use a result which was conjectured by Serge Lang and proved by Ihara, Serre and Tate (see [L1] and [L2]), which states that the intersection of an irreducible curve in $\mathbb{C}^* \times \mathbb{C}^*$ with the roots of unity $\mu_\infty \times \mu_\infty$ is finite, unless the curve is of the form $X^n Y^m - \zeta = 0$ or $X^m - \zeta Y^n = 0$ with $\zeta \in \mu_\infty$, that is, unless the curve is the translate of an algebraic subgroup by a torsion point of $\mathbb{C}^* \times \mathbb{C}^*$. Applying this result to the rational curve $\{(f(t), g(t)) : t \in \mathbb{C}\}$, we conclude that only for finitely many t's both $f(t)$ and $g(t)$ are roots of unity when f and g are multiplicatively independent.

Thus by Lang's theorem there is only a finite set of points $S \subset \mathbb{C}$ such that for any $s \in S$ both $f(s)$ and $g(s)$ are roots of unity. So $\gcd(f^k - 1, g^k - 1)$ can only have linear factors from $\{t - s \mid s \in S\}$. Write

$$f^k - 1 = \prod_{j=0}^{k-1} (f - \zeta_k^j).$$

Any two factors on the right side are coprime, so $t - s$ can divide at most one of them with multiplicity at most $\deg(f)$, and a similar statement can be said for g. Therefore the required polynomial h can be chosen as

$$h(t) = \prod_{s \in S} (t - s)^{\min(\deg(f), \deg(g))}.$$

For the second part of Theorem 1, let $s \in S$ and let d_s be the least positive integer such that
$$t - s \mid \gcd(f(t)^{d_s} - 1, g(t)^{d_s} - 1).$$
Then $d_s > 1$ because $\gcd(f-1, g-1) = 1$, and clearly for $k \notin d_s\mathbb{N}$,
$$t - s \nmid \gcd(f(t)^k - 1, g(t)^k - 1).$$
Then $\bigcup_{s\in S} d_s\mathbb{N}$ is the required finite union of proper arithmetic progressions outside which $\gcd(f^k - 1, g^k - 1) = 1$. ∎

Note that Theorem 3 implies Theorem 1. We have chosen to give the proof of Theorem 1 separately to illustrate the ideas in a simple context.

3. 2×2 **matrices.** Let $A \in \mathrm{SL}_2(\mathbb{Z})$ be a 2×2 unimodular matrix which is *hyperbolic*, that is, A has two distinct real eigenvalues. We show:

PROPOSITION 4. *Let* $A \in \mathrm{SL}_2(\mathbb{Z})$ *be a hyperbolic matrix with eigenvalues* $\varepsilon, \varepsilon^{-1}$, *where* $|\varepsilon| > 1$. *Then* $\gcd(A^k - I) \gg |\varepsilon|^{k/2}$.

Proof [1]. Let K be the real quadratic field $\mathbb{Q}(\varepsilon)$ and $\mathcal{O}_K$ its ring of integers. We may diagonalize the matrix A over K, that is, write $A = P\left(\begin{smallmatrix}\varepsilon & 0\\ 0 & \varepsilon^{-1}\end{smallmatrix}\right)P^{-1}$ with P a 2×2 matrix having entries in K. Since P is only determined up to a scalar multiple, we may, after multiplying P by an algebraic integer of $\mathcal{O}_K$, assume that P has entries in $\mathcal{O}_K$. Then $P^{-1} = (1/\det(P))P^{\mathrm{ad}}$ where P^{ad} also has entries in $\mathcal{O}_K$. Thus we have
$$A^k - I = \frac{1}{\det(P)}\, P \begin{pmatrix} \varepsilon^k - 1 & 0 \\ 0 & \varepsilon^{-k} - 1 \end{pmatrix} P^{\mathrm{ad}}.$$

The entries of $A^k - I$ are thus $\mathcal{O}_K$-linear combinations of $(\varepsilon^k - 1)/\det(P)$ and of $(\varepsilon^{-k} - 1)/\det(P)$. We now note that
$$\varepsilon^{-k} - 1 = -\varepsilon^{-k}(\varepsilon^k - 1)$$
and thus the entries of $A^k - I$ are all $\mathcal{O}_K$-multiples of $(\varepsilon^k - 1)/\det(P)$. In particular, $\gcd(A^k - I)$, which is a $\mathbb{Z}$-linear combination of the entries of $A^k - I$, can be written as
$$\gcd(A^k - I) = \frac{\varepsilon^k - 1}{\det(P)}\, \gamma_k$$
with $\gamma_k \in \mathcal{O}_K$.

Now taking norms from K to $\mathbb{Q}$ we see that
$$|\gcd(A^k - I)|^2 = \frac{|\mathcal{N}(\varepsilon^k - 1)|}{|\mathcal{N}(\det P)|}\, |\mathcal{N}(\gamma_k)|.$$

---

[1] We thank the referee for suggesting this proof, which replaces our original, more complicated, version.

Since $\gamma_k \neq 0$, we have $|\mathcal{N}(\gamma_k)| \geq 1$ and thus

$$|\gcd(A^k - I)|^2 \geq \frac{|\mathcal{N}(\varepsilon^k - 1)|}{|\mathcal{N}(\det P)|} \gg \varepsilon^k,$$

which gives $|\gcd(A^k - I)| \gg \varepsilon^{k/2}$. ■

A special case of this proposition appeared as a problem in the 54th W. L. Putnam Mathematical Competition, 1994 (see [An, pp. 82, 242]).

4. Cyclotomic fields. A standard construction of unimodular matrices is to take a unit u of norm one in a number field K and let it act by multiplication on the ring of integers $\mathcal{O}_K$ of K. This gives a linear map, and a choice of integer basis of $\mathcal{O}_K$ gives us an integer matrix whose determinant equals the norm of u and is thus unimodular. We employ this method for the case when u is a nonreal unit to give a construction of matrices A with the property that A^k is primitive infinitely often.

We recall some basic facts on units in a cyclotomic field. Let $p > 3$ be a prime, ζ_p a primitive pth root of unity, and $K = \mathbb{Q}(\zeta_p)$ the cyclotomic extension of the rationals. It is a field of degree $p-1$. The ring of integers $\mathcal{O}_K$ of this field is $\mathbb{Z}[\zeta_p]$. Since K is purely imaginary, it follows that the norm function is positive, and the norm of a unit u is always 1. Also note that the structure of the unit group E_p of $\mathcal{O}_K$ is

$$E_p = W_p E_p^+, \tag{3}$$

where W_p are the roots of unity in K and E_p^+ is the group of the real units in $\mathcal{O}_K$. A proof of this fact can be found, for example, in [L3, Theorem 4.1].

4.1. *Proof of Theorem 2.* We now prove Theorem 2, that is, show that if $u \in E_p \backslash E_p^+$ is a nonreal unit and $k \not\equiv 0 \bmod p$ then the matrix corresponding to u^k is primitive.

The method we will use is that if we choose a basis $\omega_0 = 1, \omega_1, \ldots, \omega_{p-2}$ of $\mathbb{Z}[\zeta_p]$ and take a unit U in $\mathbb{Z}[\zeta_p]$, then we get a matrix $A(U) = (a_{i,j})$ whose entries are determined by

$$U\omega_i = \sum_{j=0}^{p-2} a_{j,i}\omega_j.$$

In particular if we find that in the expansion of

$$U = U\omega_0 = \sum_{j=0}^{p-2} a_{j,0}\omega_j$$

we have an index $j \neq 0$ so that $a_{j,0} = a_{0,0}$, then in the matrix $A(U) - I$ corresponding to $U - 1$, the first column will contain the entries $a_{0,0} - 1$ and $a_{j,0} = a_{0,0}$ which are clearly coprime, and thus the matrix $A(U)$ is *primitive.*

Another option is to have $a_{0,0} = 0$, in which case in the matrix of $U - 1$, the $(0,0)$ entry is -1, and thus again $A(U)$ is primitive. We will apply this method to the case that $U = u^k$ is a power of a nonreal unit u and $k \not\equiv 0 \bmod p$.

Let $u \in E_p \backslash E_p^+$ be a nonreal unit. By (3), we can write

$$u = \zeta_p^x u^+,$$

where u^+ is a *real* unit and x is an integer not congruent to 0 mod p. Therefore, $u^k = \zeta_p^{xk}(u^+)^k$ and $\zeta_p^{-xk}u^k = (u^+)^k$ is real. Hence it can be represented as an integer combination of $\zeta_p, \zeta_p^2, \ldots, \zeta_p^{p-1}$ as follows:

$$\zeta_p^{-xk}u^k = \sum_{j=1}^{p-1} \alpha_j \zeta_p^j,$$

where $\alpha_j = \alpha_{p-j}$ for each j. For convenience we will set $\alpha_0 := 0$.

Multiplying by ζ_p^{xk}, we find

$$u^k = \sum_{j=0}^{p-1} \alpha_j \zeta_p^{j+xk}$$

and changing the summation variable,

$$u^k = \sum_{i=0}^{p-1} \alpha_{i-xk} \zeta_p^i,$$

where the index of α is calculated mod p. Using the relation

$$\zeta_p^{p-1} = -1 - \zeta_p - \ldots - \zeta_p^{p-2}$$

we find that in terms of the integer basis $\omega_j = \zeta_p^j$, $j = 0, \ldots, p-2$, we have

$$u^k = \sum_{i=0}^{p-2} (\alpha_{i-xk} - \alpha_{p-1-xk})\omega_i.$$

If $k \not\equiv 0 \bmod p$ then $2xk \not\equiv 0 \bmod p$ since $x \not\equiv 0 \bmod p$. If $2xk \not\equiv -1 \bmod p$ then the coefficients of ω_0 and ω_{2xk} are equal. Therefore u^k is primitive. If $2xk$ is congruent to $-1 \bmod p$, then the coefficient of ω_0 vanishes and thus in this case as well, u^k is primitive.

Thus we have found that if $k \not\equiv 0 \bmod p$, the matrix corresponding to u^k is primitive. ∎

Note that by virtue of (3), the eigenvalues of $A(u)$ come in complex conjugate pairs whose ratios are pth roots of unity. This is somewhat similar to the trivial scalar example described in the introduction, namely $b = \pm a$.

5. Proof of Theorem 3. We extend the idea of the proof of Theorem 1 to cover the matrix case. We first show that there is only a finite set S of points $s \in \mathbb{C}$ such that $t - s$ divides $\gcd(A^k - I)$ for some k.

Let M be a matrix such that MAM^{-1} is in Jordan form. The elements of M are meromorphic functions on the Riemann surface R corresponding to some finite extension of $\mathbb{C}(t)$. Denote by $\mathrm{pr} : R \to \mathbb{P}^1$ the associated projection of R to the projective line. Let S_0 be the finite collection of poles of these functions.

Assume first that A is not diagonalizable over the algebraic closure of $\mathbb{C}(t)$. Thus for any $t_0 \in R \setminus S_0$, $A(t_0)$ is not diagonalizable, and therefore $A(t_0)^k - I \neq 0$ for all k (recall that a matrix of finite order ($A^m = I$) is automatically diagonalizable), in other words, $t - t_0$ does not divide $\gcd(A^k - I)$. Thus only the finitely many linear forms $t - s$, where $s \in \mathrm{pr}(S_0)$ is the projection of some point in S_0, can divide $\gcd(A^k - I)$.

We denote by $\lambda_i(t)$ the eigenvalues of A which are multivalued functions of t, that is, meromorphic functions on the Riemann surface. Assume now that λ_1 and λ_2 are multiplicatively independent, and that A is diagonalizable. Suppose that $(t - t_0) \mid \gcd(A^k - I)$ for some $k > 1$ and $t_0 \in R \setminus S_0$. Then $A^k - I$ evaluated at t_0 is the zero matrix, and also

$$M(t_0)(A(t_0)^k - I)M(t_0)^{-1} = 0,$$

and we deduce that

$$\lambda_1(t_0)^k - 1 = \lambda_2(t_0)^k - 1 = 0.$$

In particular, $\lambda_1(t_0)$ and $\lambda_2(t_0)$ are roots of unity. Thus, we have reduced our task to proving that λ_1 and λ_2 can be simultaneous roots of unity only at a finite set of points.

To prove this, we want to use Lang's theorem for the curve in $\mathbb{C}^2$ parameterized by $(\lambda_1(t), \lambda_2(t))$. Denote by Y the Zariski closure of the image of the map $(\lambda_1, \lambda_2) : R \setminus S_0 \to \mathbb{C}^2$. Then Y is an irreducible algebraic curve in $\mathbb{C}^2$. If Y is of dimension 0, then it is a point, so $\lambda_1(t)$ and $\lambda_2(t)$ are constants, and since they are multiplicatively independent none of them can be a root of unity. Otherwise, we may apply Lang's theorem for this curve to conclude that unless the curve Y is of the form $F^m - \zeta G^n = 0$ or $F^m G^n = \zeta$ with ζ a root of unity (which is not the case when λ_1 and λ_2 are multiplicatively independent), it has only finitely many torsion points. In other words, there can only be finitely many points of the form (ζ_1, ζ_2) on Y, where ζ_1 and ζ_2 are roots of unity.

We now prove that there is a polynomial h such that $\gcd(A^k - I)$ divides h for all k. Since there is a finite set S of possible zeros of $\gcd(A^k - I)$, it suffices to show that the multiplicity of a zero of $\gcd(A^k - I)$ is bounded.

Write $B = MAM^{-1}$, so B is in Jordan form. Denote by $v_{t_0}(f)$ the multiplicity of the zero of f at $t_0 \in R$. So clearly, for any $t_0 \in R$ there exists

$c(t_0) \in \mathbb{N}$ such that

$$v_{t_0}(\gcd(A^k - I)) \le c(t_0) + v_{t_0}(\gcd(B^k - I)),$$

and for all t_0 outside the finite set S_0 of poles of entries of M, $c(t_0) = 0$. So it suffices to prove that $v_{t_0}(\gcd(B^k - I))$ is bounded. Clearly,

$$\gcd(B^k - I) \mid \det(B^k - I) = \prod_{j=0}^{k-1} \det(B - \zeta_k^j I),$$

where ζ_k is a primitive kth root of unity. Denoting the diagonal elements of $B - I$ by $b_1, \ldots, b_r$, we see that

$$\det(B^k - I) = \prod_{d=1}^{r} \prod_{j=0}^{k-1} (b_d - \zeta_k^j).$$

Because a meromorphic function on a Riemann surface has a finite degree, reasoning as in the proof of Theorem 1 we see that for any $t_0 \in R$, $v_{t_0}(\prod_{j=1}^{k}(b_d - \zeta_k^j))$ is bounded, for all k. Therefore $v_{t_0}(\det(B^k - I))$ is bounded for all k.

Now assume in addition that A is primitive: $\gcd(A - I) = 1$. For any $s \in S$, the set of k's such that $A(s)^k = I$, i.e. $(t - s) \mid \gcd(A^k - I)$, is an arithmetic progression $d_s\mathbb{Z}$ which is proper since it does not contain 1. Therefore the set of k's with $\gcd(A^k - I) \ne 1$ is a finite union of proper arithmetic progressions, and hence for k outside this union, we have $\gcd(A^k - I) = 1$. ∎

References

[Ai] N. Ailon, *Primitive powers of matrices and related problems*, M.Sc. thesis, Tel Aviv Univ., 2001.

[An] T. Andreescu and R. Gelca, *Mathematical Olympiad Challenges*, Birkhäuser Boston, Boston, MA, 2000.

[BCZ] Y. Bugeaud, P. Corvaja and U. Zannier, *An upper bound for the G.C.D. of $a^n - 1$ and $b^n - 1$*, Math. Z. 243 (2003), 79–84.

[L1] S. Lang, *Division points on curves*, Ann. Mat. Pura Appl. (4) 70 (1965), 229–234.

[L2] —, *Fundamentals of Diophantine Geometry*, Springer, 1983, 200–207.

[L3] —, *Cyclotomic Fields*, Springer, 1978, 79–82.

Department of Computer Science
Princeton University
Princeton, NJ 08544, U.S.A.
E-mail: nailon@princeton.edu

Raymond and Beverly Sackler School
of Mathematical Sciences
Tel Aviv University
Tel Aviv 69978, Israel
E-mail: rudnick@post.tau.ac.il

Received on 4.7.2002
and in revised form on 15.12.2002 (4323)