TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 273, Number 1, September 1982

THE CATALAN EQUATION OVER FUNCTION FIELDS

BY

JOSEPH H. SILVERMAN

ABSTRACT. Let K be the function field of a projective variety. Fix $a, b, c \in K^*$. We show that if $\max\{m, n\}$ is sufficiently large, then the Catalan equation $ax^m + by^n = c$ has no nonconstant solutions $x, y \in K$.

The Cassels-Catalan conjecture states that for fixed nonzero integers a, b, c, the equation $ax^m + by^n = c$ has only finitely many solutions in integers x, y, m, n satisfying $m \geqslant 3$ and n, $|x|$, $|y| \geqslant 2$. At present, the only known case of this conjecture is $a = -b = c = 1$, due to Tijdeman **[6]**. In this paper we prove a strengthened version of the Cassels-Catalan conjecture in the case that the number field $\mathbf{Q}$ is replaced by an arbitrary function field. The proof uses only elementary algebraic geometry, the principal tool being the Riemann-Hurwitz formula.

THEOREM. *Let k be a field of characteristic p (possibly with $p = 0$), and let K/k be the function field of a nonsingular projective variety. Fix $a, b, c \in K^*$.*

Then there are only finitely many pairs of integers $m, n \geqslant 2$ (prime to p if $p \neq 0$) for which the Cassels-Catalan equation $ax^m + by^n = c$ has even a single nonconstant solution $x, y \in K$, $x, y \notin k$.

Further, for any particular pair m, n as above, there will be only finitely many solutions $x, y \in K$ unless either:

(i) *a/c is an mth power and b/c is an nth power in K, in which case there may be infinitely many solutions $(x, y) = (\alpha(a/c)^{1/m}, \beta(b/c)^{1/n})$ with $\alpha, \beta \in k$ satisfying $\alpha^m + \beta^n = 1$; or*

(ii) *$(m, n) \in \{(2,2), (2,3), (3,2), (2,4), (4,2), (3,3)\}$, in which case the Cassels-Catalan equation defines a curve of genus 0 or 1 over K.*

PROOF. We first note that taking "generic" hyperplane sections of the variety whose function field is K, we are reduced by Bertini's theorem to the case that $K = k(C)$ is the function field of a nonsingular projective curve C. (See, e.g., **[5]** for the details of this standard reduction.) Second, we may replace k by its algebraic closure, since at worst this will create extra solutions. Third, dividing the equation by c and replacing a and b by a/c and b/c, we may assume that our equation is

$$(*) \qquad ax^m + by^n = 1.$$

Let $D = \max\{\deg(a), \deg(b)\}$. (Note a and b are now functions on the curve C, so have degrees.) By symmetry, we may assume that $m \geqslant n$.

Received by the editors May 28, 1981 and, in revised form, July 24, 1981.
1980 *Mathematics Subject Classification.* Primary 10B15; Secondary 10J06, 14J25.

0002-9947/81/0000-0631/$02.25

Consider the desingularization of the covering of C given by the equations $u^m = a$, $v^n = b$. Since this is a composition of cyclic coverings (note that k contains both an mth and an nth root of unity), it will consist of a union of isomorphic curves. Let C' be any one of these curves, and f: $C' \to C$ the natural map. We note for future reference that the degree of f divides mn; in particular, $\deg(f)$ is prime to p if $p > 0$.

Let V/C be the projective surface given by the equation $ax^m + by^n z^{m-n} = z^m$. Let $V' = V \times_C C'$ be base extension of V by C'. If V_0 is the projective *curve* $X^m + Y^n Z^{m-n} = Z^m$ then we have a natural map

$$V' \to V_0,$$
$$[x, y, z] \to [ux, vy, z].$$

Further let V_0^* be the desingularization of V_0 and h: $V_0^* \to V_0$ the natural map.

Now suppose that we are given functions $x, y \in k(C)$ which satisfy equation $(*)$. Then $P = [s, y, 1]$ gives a section P: $C \to V$. This extends to a section $C' \to V'$, and composed with the map $V' \to V_0$ from above, gives a map of curves $C' \to V_0$. But C' is nonsingular, so it factors through the normalization map h to yield a map ϕ: $C' \to V_0^*$. All of this is summarized in the following diagram.

$$\begin{tikzcd}
V_0^* \arrow[r, "h"] & V_0 & V' \arrow[l] \arrow[r] \arrow[d] & V \arrow[d, shift left] \\
& & C' \arrow[ull, "\phi"] \arrow[ul, "{[ux, vy, 1]}"'] \arrow[r, "f"] & C \arrow[u, dashed, shift left, "P"']
\end{tikzcd}$$

Assume for now that ϕ is surjective. (The other case is dealt with later.) We apply the Riemann-Hurwitz formula twice, once to ϕ and once to f. Note that since $\deg(f)$ is prime to $p = \operatorname{char}(k)$, there is no wild ramification in f even if $p > 0$. (See, e.g., [**3**, Chapter IV. 2] for basic facts about the Riemann-Hurwitz formula.)

$$(**) \qquad \begin{aligned} [2g(V_0^*) - 2]\deg(\phi) &\leq 2g(C') - 2 \\ &= [2g(C) - 2]\deg(f) + \sum_{t \in C} [e_t(f) - 1], \end{aligned}$$

where $e_t(f) = \deg(f) - \#f^{-1}(t) + 1$. Now $e_t(f) \leq \deg(f)$, and the only points $t \in C$ at which f can possibly ramify are zeros and poles of a and b. Thus $e_t(f) = 1$ except possibly at $4D$ points of C, hence

$$\sum_{t \in C} [e_t(f) - 1] \leq 4D \deg(f).$$

It is an elementary exercise to resolve the singularity of V_0 at $[0, 1, 0]$ (which is singular if and only if $m - n > 1$) and compute the Euler characteristic of V_0^*. One finds

$$2g(V_0^*) - 2 = mn - m - n - (m, n),$$

where (m, n) is the greatest common divisor of m and n.

Next we express $\deg(\phi)$ in terms of x and y. Consider the following commutative diagram.

$$\begin{tikzcd} V_0^* \arrow[r, "\text{degree } 1", "h"'] \arrow[d, dash, "\phi"'] & V_0 \arrow[rr, "\text{degree } mn"] & & \mathbf{P}^1 \\ & {[X, Y, Z]} \arrow[r, mapsto] & {[X^m, Z^m]} & \\ C' \arrow[rrr, "f"'] & & & C \arrow[uu, "{[ax^m, 1]}"'] \end{tikzcd}$$

From this we read off

$$\deg(\phi) = \frac{\deg(ax^m)\deg(f)}{mn} \geqslant \frac{(\deg(x^m) - D)\deg(f)}{mn}.$$

Using a similar diagram we obtain

$$\deg(\phi) = \frac{\deg(by^n)\deg(f)}{mn} \geqslant \frac{(\deg(y^n) - D)\deg(f)}{mn}.$$

Now putting all of these computations into the inequality (**) and dividing by $\deg(f)$, we obtain

$$\begin{aligned}\left(5 - m^{-1} - n^{-1} - [m, n]^{-1}\right)D + 2g(C) - 2 \\ \geqslant \left(1 - m^{-1} - n^{-1} - [m, n]^{-1}\right)\max\{\deg(x^m), \deg(y^n)\},\end{aligned}$$

where $[m, n]$ is the least common multiple of m and n. In particular we obtain the fundamental inequality

$$(+) \quad 5D + 2g(C) - 2 \geqslant \left(1 - m^{-1} - n^{-1} - [m, n]^{-1}\right) \max\{\deg(x^m), \deg(y^n)\}.$$

We now check the case that ϕ is a constant map. This will mean that ux and vy are constant functions on C', so raising to the mth (respectively nth) power we find that ax^m and by^n are in the constant field k. Hence, since $ab \neq 0$, $\deg(x^m) \leqslant \deg(a) \leqslant D$ and $\deg(y^n) \leqslant \deg(b) \leqslant D$. These are stronger than the inequality $(+)$ except in the one trivial case $D = g(C) = 0$, in which case they imply that $\deg(x) = \deg(y) = 0$, while $(+)$ would yield the impossibility $-2 \geqslant 0$. In any case, $(+)$ holds whenever x and y are not constant functions.

It is easy to check that if $m, n \geqslant 2$ are integers other than the six pairs listed in (ii) of the theorem, then

$$1 - 1/m - 1/n - 1/[m, n] \geqslant 1/6$$

(attaining this minumum for $(m, n) = (6, 2)$). Now since our solution $x, y \in k(C)$ to $(*)$ was assumed to be nonconstant, we have $\deg(x)$, $\deg(y) \geqslant 1$, so the inequality $(+)$ yields

$$5D + 2g(C) - 2 \geqslant \tfrac{1}{6}\max\{m, n\}.$$

This bounds the possibilities for m and n. Then, for any particular m and n,

$$5D + 2g(C) - 2 \geqslant \tfrac{1}{6}\max\{m\deg(x), n\deg(y)\}$$

bounds $\deg(x)$ and $\deg(y)$. But a curve of genus at least 2 over a function field K can have only finitely many points of bounded degree unless the curve is birational, over K, to a curve defined over the field of constants k. Further, all but finitely many of thoses points will come from k-valued points on the new curve. (See [**5**, Corollary on p. 42]. Note that the result is true also for $\operatorname{char}(k) = p > 0$, since we have taken k to be algebraically closed. In general, one must take a purely inseparable extension of k.) One easily checks that the only way for the Cassels-Catalan equation $(*)$ to reduce to an equation over k is for a to be a mth power and b to be an nth power, which is the case covered by (i) of the theorem.

The inequality $(+)$ derived in the course of proving the above theorem is of independent interest, since it gives effective bounds for m, n, $\deg(x)$, $\deg(y)$ in the case that K is the function field of a curve. We therefore state it separately.

THEOREM. *Let k be a field and C/k a nonsingular projective curve with function field $k(C)$. Fix two functions a, $b \in k(C)^*$. Suppose m, $n \geqslant 2$ are integers [prime to* $\operatorname{char}(k)$ *if* $\operatorname{char}(k) \neq 0$] *and x, $y \in k(C)$ are functions satisfying $ax^m + by^n = 1$. Then*

$$\begin{aligned} &5\max\{\deg(a), \deg(b)\} + 2g(C) - 2 \\ &\qquad \geqslant \left(1 - m^{-1} - n^{-1} - [m, n]^{-1}\right)\max\{\deg(x^m), \deg(y^n)\}. \end{aligned}$$

(*except in the trivial case* $g = \deg(x) = \deg(y) = \deg(a) = \deg(b) = 0$).

It is likely that the coefficient 5 in the left-hand side of this inequality can be improved, and one might ask what the best possible result is. In the special case that $g = 0$ (so $k(C) = k(t)$ is a rational function field), if one restricts x and y to be polynomials, then Davenport has used very different arguments to obtain a similar inequality [**1**]. If $g = 0$, $a = b = 1$, and $m = n$ (i.e. the Fermat equation over $k(t)$), one can show the nonexistence of nonconstant solutions in $k(t)$ by an easy descent argument [**2**]. Both of these proofs, however, use the unique factorization of $k[t]$, so do not readily generalize to other functions fields.

Returning to the case of number fields, one is tempted to conjecture that an inequality of this sort should hold (with $2g(C) - 2$ replaced by some function involving the degree of the number field over $\mathbf{Q}$). From this the strengthened version of the Cassels-Catalan conjecture would follow formally as in the above proof. (Notice a corollary would be Fermat's last theorem for sufficiently large exponents!) A similar inequality, but only for integral rather than rational solutions, has been proposed by Lang and Waldschmidt [**4**, p. 213]. They also show how it would follow from a certain Baker-style diophantine inequality, but this unfortunately seems well beyond current techniques.

ACKNOWLEDGEMENT. I would like to thank Barry Mazur for his encouragement and helpful suggestions.

REFERENCES

1. H. Davenport, *On* $f^3(t) - g^2(t)$, Norske Vid. Selsk. Forh. (Trondheim) **38** (1965), 86–87.

2. N. Greenleaf, *On Fermat's equation in* $C(t)$, Amer. Math. Monthly **76** (1969), 808–809.

3. R. Hartshorne, *Algebraic geometry*, Graduate Texts in Math., vol. 52, Springer-Verlag, Berlin and New York, 1977.

4. S. Lang, *Elliptic curves: Diophantine analysis*, Grundlehren der Math. Wissenschaften, vol. 231, Springer-Verlag, Berlin and New York, 1978.

5. ______, *Integral points on curves*, Inst. Hautes Études Sci. Publ. Math. **6** (1960), 27–43.

6. R. Tijdeman, *On the equation of Catalan*, Acta Arith. **29** (1976), 197–209.

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS 02138