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 $m > 3$ and $n, |x|, |y| \geq 2$. 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 $\mathbb{Q}$ 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 $m, n \geq 2$ (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 $m$th power and $b/c$ is an $n$th 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

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

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 $m \geq n$. 

Received by the editors May 28, 1981 and, in revised form, July 24, 1981. 
1980 Mathematics Subject Classification. Primary 10B15; Secondary 10J06, 14J25.
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 \( m \)th and an \( n \)th 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 \( V' \to V_0 \), 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 \( \phi: C' \to V_0^* \). All of this is summarized in the following diagram.

\[
\begin{array}{ccc}
V_0^* & \xrightarrow{h} & V_0 \\
\phi \downarrow & & \downarrow f \\
C' & \xrightarrow{P} & C
\end{array}
\]

Assume for now that \( \phi \) is surjective. (The other case is dealt with later.) We apply the Riemann-Hurwitz formula twice, once to \( \phi \) and once to \( f \). Note that since \( \deg(f) \) is prime to \( p = \text{char}(k) \), 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.)

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

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.

\[
\begin{array}{cccccc}
V^* & \xrightarrow{h} & V_0 & \xrightarrow{[X, Y, Z]} & \mathbb{P}^1 \\
\phi & & \downarrow & & \\
C' & \xrightarrow{f} & C
\end{array}
\]

From this we read off

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

Using a similar diagram we obtain

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

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

\[
(5 - m^{-1} - n^{-1} - [m, n]^{-1})D + 2g(C) - 2 \\
\geq (1 - m^{-1} - n^{-1} - [m, n]^{-1}) \max\{\deg(x^m), \deg(y^n)\},
\]

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

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

We now check the case that $\phi$ is a constant map. This will mean that $ux$ and $vy$ are constant functions on $C'$, so raising to the $m$th (respectively $n$th) power we find that $ax^m$ and $by^n$ are in the constant field $k$. Hence, since $ab \neq 0$, $\deg(x^m) \leq \deg(a) \leq D$ and $\deg(y^n) \leq \deg(b) \leq D$. 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 $-2 \geq 0$. In any case, (+) holds whenever $x$ and $y$ are not constant functions.

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

\[
1 - 1/m - 1/n - 1/[m, n] \geq 1/6
\]

(attaining this minimum for $(m, n) = (6, 2)$). Now since our solution $x, y \in k(C)$ to (9) was assumed to be nonconstant, we have $\deg(x), \deg(y) \geq 1$, so the inequality (+) yields

\[
5D + 2g(C) - 2 \geq \frac{1}{6} \max\{m, n\}.
\]
This bounds the possibilities for $m$ and $n$. Then, for any particular $m$ and $n$,

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

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 those 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 $\text{char}(k) = p > 0$, 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 $m$th power and $b$ to be an $n$th 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 $m$, $n \geq 2$ are integers [prime to $\text{char}(k)$ if $\text{char}(k) \neq 0$] and $x$, $y \in k(C)$ are functions satisfying $am^x + bn^y = 1$. Then

$$5\max \{ \deg(a), \deg(b) \} + 2g(C) - 2$$

$$\geq \left(1 - m^{-1} - n^{-1} - [m, n]^{-1}\right)\max \{ \deg(x^m), \deg(y^n) \}.$$  

(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 $Q$). 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.

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MASSACHUSETTS 02138