JOURNAL OF NUMBER THEORY 47, 131–137 (1994)

A Small Contribution to Catalan's Equation

A. M. W. GLASS, DAVID B. MERONK, TAIHEI OKADA, AND RAY P. STEINER

Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, Ohio 43403-0221

Communicated by Alan C. Woods

Received January 2, 1992; revised March 19, 1992

DEDICATED TO THE MEMORY OF PROFESSOR HANS ZASSENHAUS

Using recent results on linear forms in logarithms of algebraic numbers, we prove that any solution of the equation $x^p - y^q = \varepsilon$, where $\varepsilon = \pm 1$, p and q are odd primes, and $p > q$ satisfies $p < 3.42 \cdot 10^{28}$ and $q < 5.6 \cdot 10^{19}$. We also combine our work with some results of Altonen and Inkeri to determine the six cases with $q \leqslant 37$ for which this equation may have solutions.

The purpose of this note is to show that if Catalan's equation has any non-trivial integer solutions other than $3^2 - 2^3 = 1$, then it has such solutions with relatively small exponents. Specifically:

THEOREM. *If* $x^m - y^n = 1$ *has a solution for* (x, y, m, n) *in the set of positive rational integers greater than* 1 *other than* $(3, 2, 2, 3)$, *then it has such a solution with* $\max\{m, n\} < 3.42 \times 10^{28}$ *and* $\min\{m, n\} < 5.6 \times 10^{19}$.

This improves Langevin's constant [6] of over 10^{100} for $\max\{m, n\}$. There are explicit bounds on x and y as a consequence of Alan Baker's work on hyperelliptic functions (see [3] or [10, Chap. 6]).

In addition, we remove many small values of $\min\{m, n\}$.

In his superb work [11], Tijdeman showed that Catalan's equation $x^m - y^n = 1$ has only a finite number of solutions in the rational integers for x, y, m, and n greater than 1. His proof relied heavily on Alan Baker's brilliant analysis of linear forms in the logarithms of algebraic numbers [2]. However, Tijdeman used only the qualitative aspect of Baker's results; the quantitative aspect does give explicit bounds on x, y, m, and n, as was noted by Tijdeman. In this note, we wish to use the new improvements on the constants for linear forms in the logarithms of algebraic numbers [12, 7] to obtain the theorem. Our approach is to effectivize Tijdeman's proof as given in [10, Chap. 12, pp. 205ff.].

0022-314X/94 $6.00

Proof of Theorem. If there exists any solution in the rational integers greater than 1 other than $3^2 - 2^3 = 1$, there exists one with prime exponents. Indeed, if we let $\varepsilon = \pm 1$, p and q be primes, and $p > q$, we need only show that if $x^p - y^q = \varepsilon$ has a solution in the positive rational integers, then it has one such with $p < 3.42 \times 10^{28}$ and $q < 5.6 \times 10^{19}$, other than $3^2 - 2^3 = 1$.

By [8], we may assume that $q \geqslant 5$. Moreover, by [5], if $x^p - y^q = \varepsilon$, then $p \mid y$ and $q \mid x$. Therefore,

$$0 \equiv x^p = y^q + \varepsilon \equiv y + \varepsilon \qquad (\operatorname{mod} q)$$

and similarly, $x - \varepsilon \equiv 0 \pmod p$. Now, following Shorey and Tijdeman [10]—and using the same numbering—we have

$$y + \varepsilon = q^{-1}s^p \tag{13}$$

and

$$x - \varepsilon = p^{-1}r^q \tag{14}$$

for some positive rational integers r, s with $p \mid r$ and $q \mid s$ (see [10, p. 205]). Hence $p^{-1}r^q \geqslant p^{q-1} \geqslant 7^{q-1}$ and $q^{-1}s^p \geqslant 5^{p-1}$, since $r \geqslant p \geqslant 7$ and $s \geqslant q \geqslant 5$. By (13) and (14) we have

$$(p^{-1}r^q + \varepsilon)^p - (q^{-1}s^p - \varepsilon)^q = \varepsilon. \tag{15}$$

Moreover, $r^{pq} \geqslant (p^{-1}r^q + 1)^p + 1 \geqslant x^p + 1 \geqslant y^q \geqslant (q^{-1}s^p - 1)^q \geqslant s^{pq}/(2q)^q$. Similarly $s^{pq} \geqslant y^q + 1 \geqslant x^p \geqslant r^{pq}/(2p)^p$. Thus

$$\begin{aligned} s &\leqslant r(2q)^{1/q}, \\ r &\leqslant s(2p)^{1/q}. \end{aligned} \tag{16}$$

We wish to establish that

$$q \leqslant e^{32.91}(\log p)^3, \tag{17}$$

an improvement on [10] with an explicit constant. Since $q \geqslant 5$ and $p \mid r$, it follows that $p^{-1}r^q \geqslant p^{q-1} \geqslant p^4$. Now, by (13) and (14),

$$\left|\frac{px}{r^q} - 1\right| = \frac{p}{r^q}, \qquad \left|\frac{qy}{s^p} - 1\right| = \frac{q}{s^p} \qquad \text{and} \qquad \left|\frac{y^q}{x^p} - 1\right| = \frac{1}{x^p}.$$

Since $|\log(1 + \alpha)| \leqslant 2\,|\alpha|$ whenever $|\alpha| \leqslant \frac{1}{2}$, we get, by (16),

$$|p \log(r^q/p) - p \log x| \leqslant 2p^2/r^q \tag{19}$$

$$|p \log x - q \log y| \leqslant 2/x^p \leqslant 2p/r^q \tag{20}$$

$$|q \log y - q \log(s^p/q)| \leqslant 2q^2/s^p \leqslant \frac{2q^2}{s^2} \cdot \frac{1}{s^q} \leqslant 2\left(\frac{2p}{r^q}\right) \leqslant \frac{4p}{r^q}. \tag{21}$$

Let

$$\Lambda_1 = |p \log(r^q/p) - q \log(s^p/q)|.$$

Then $\Lambda_1 \leqslant 4p^2/r^q$ since $p \geqslant 7$.

Note that

$$\Lambda_1 = |pq \log(r/s) + q \log q - p \log p|.$$

If $x^p - y^q = \varepsilon$, then exactly one of x and y is odd. Thus exactly one of $x - \varepsilon$ and $y + \varepsilon$ is even. Since p and q are odd, r or s is even (and not both). Consequently $\{\log q, \log p, \log(r/s)\}$ is a linearly independent set over Q, so $\Lambda_1 \neq 0$.

We now assume that $p \geqslant 10^{27}$ and apply [12, Theorem 2.18; Sect. 9, Table 2]. In the notation given there, $A_1 = q$, $A_2 = p$, and $A_3 = 2r$ (by (16) above). Let $E = 5 \leqslant q, p, 2r$; for

$$E \leqslant \frac{3}{f}\left(1 + 1 + \frac{|\log(r/s)|}{\log 2r}\right)^{-1},$$

it suffices that

$$E \leqslant \frac{3}{f}\left(1 + 1 + \frac{\log(2p)^{1/q}}{\log 2r}\right)^{-1}$$

by (16). Since $q \geqslant 5$ and $r \geqslant p \geqslant 10^{27}$, we need only require that $5 \leqslant (3/f)(2.21)^{-1}$; hence we let $f = 0.27$. Now $M \leqslant pq$, $Z_0 \leqslant 10.3$, *and* $G_0 \leqslant 2 \log p$ (since $p \geqslant 10^{21}$). Thus

$$U_0 = (2 \log p)(10.3)(\log q)(\log p)(\log 2r)/(\log 5)^4,$$

and

$$\Lambda_1 \geqslant \exp\left\{-2^6 3^{14} 1950 \left(1 + \frac{1}{0.27}\right)^3 (2 \log p)(10.3)(\log q) \times (\log p)(\log 2r)/(\log 5)^4\right\}.$$

Consequently, $\Lambda_1 \geqslant \exp\{-e^{32.89}(\log p)^3 (\log 2r)\}$.

Hence $r^q \leqslant 4p^2 \exp\{e^{32.89}(\log p)^3 (\log 2r)\}$. Therefore

$$q \leqslant \frac{\log 4}{\log r} + \frac{2 \log p}{\log r} + e^{32.89}(\log p)^3 \left[1 + \frac{\log 2}{\log r}\right].$$

Now $r \geqslant p \geqslant 10^{27}$, so $q \leqslant e^{32.91}(\log p)^3$ (as desired), and $\log q \leqslant 32.91 + 3 \log\log p$.

We now apply [7] to give an explicit bound for p.
By (13) and (14),

$$(p^{-1}r^q+\varepsilon)^p-q^{-q}s^{pq}=x^p-(y+\varepsilon)^q\neq 0; \tag{24}$$

so, by (20) and (21),

$$0<|p\log x-q\log(s^p/q)|\leqslant\frac{2}{x^p}+\frac{2q^2}{s^p}.$$

Further, by (11) and (13), $x^p\geqslant y^{q-1}>2^{q/2}y>2qy>s^p$. Define

$$\Lambda_2=\left|q\log q+p\log\left(\frac{p^{-1}r^q+\varepsilon}{s^q}\right)\right|\leqslant\frac{4q^2}{s^p}. \tag{25}$$

We wish to use [7, Theorem 5.11].
First note that $p^{-1}r^q+\varepsilon=x<s^q$ (since $x^p=y^q+\varepsilon<(q(y+\varepsilon))^q=s^{pq}$).
We now show that $2e\log(s^q/x)<\log(s^q)$. Indeed,

$$\begin{aligned}2ep\log(s^q/x)&=2e\log(s^{pq}/x^p)=2e\log(q^q(y+\varepsilon)^q/(y^q+\varepsilon))\\&=2e\log(q^q)+2e\log((y+\varepsilon)^q/(y^q+\varepsilon)).\end{aligned}$$

But

$$\begin{aligned}\log((y+\varepsilon)^q/(y^q+\varepsilon))&\leqslant\log((y+1)^q/(y^q+1))=\log\left[\left(1+\frac{1}{y}\right)^q\Big/\left(1+\frac{1}{y^2}\right)\right]\\&\leqslant q\log\left(1+\frac{1}{y}\right)\leqslant\frac{2q}{y}\end{aligned}$$

since $y\geqslant 10$ (see [1], e.g.). Now

$$\frac{2q}{y}=\frac{2q}{y+\varepsilon}\cdot\frac{y+\varepsilon}{y}\leqslant\frac{2q^2}{s^p}\left(1+\frac{1}{y}\right)\leqslant\frac{2}{q^q}\left(1+\frac{1}{y}\right),$$

because $s\geqslant q$ and $p\geqslant q+2$. Since $s\geqslant q\geqslant 5$ and $y\geqslant 10$, we deduce that $2ep\log(s^q/x)\leqslant(2e+10^{-2})\log(s^q)$. As $2e+10^{-2}\leqslant 7\leqslant p$, we readily obtain our desired inequality.

We next observe that $\{\log q,\log(s^q/x)\}$ is linearly independent over Q. (Otherwise for some positive relatively prime integers m and n, $q^m=(s^q/x)^n$. Hence $x^nq^m=s^{qn}$. If a is the highest power of q that divides x and b the highest power of q that divides s, then $na+m=bqn$; so $m=n(bq-a)$. This contradicts the coprimeness of m and n unless $n=1$. If $n=1$, then $s^q=xq^m$. Thus $q^q(y+\varepsilon)^q=s^{qp}=x^pq^{mp}=(y^q+\varepsilon)\,q^{mp}$. But $q^q(y+\varepsilon)^q\leqslant$

$q^q y^q(1+1/y)^q < q^q y^q e$ (since $p \mid y$ and so $y \geqslant p > q$). Since $m \geqslant 1$, $(y^q+\varepsilon)\, q^{mp} \geqslant q^{mp}(y^q-1) \geqslant q^{q+2}y^q(1-1/y^q) > \frac{1}{2}q^{q+2}y^q e$, a contradiction.)

Let $\Lambda_2 = (q/p)\log q - \log(s^q/x) \neq 0$. We follow [7, Sect. 5.1] and define $b_1 = q$, $b_2 = p$, $B = p$, $\alpha_1 = q$, $\alpha_2 = s^q/x$, $a_1 = 2e\log q$, $a_2 = q\log s$, $f = 2e$, $\theta = 22$, and $Z = 1$. Note that $a_1 b_1 \leqslant a_2 b_2$ since $q \leqslant s$ and $2e \leqslant 7 \leqslant p$. Moreover, since $s^q > x$, $h(\alpha_2) = s^q$. Therefore $a_j \geqslant 1$, $h(\alpha_j)$, $f\,|\log\alpha_j|$ for $j = 1, 2$ by the previous paragraph. Since $q \geqslant 5$, we let

$$G' = \log\left(\frac{e}{2} + \frac{2e}{\log 5}\right) = 1 + \log\left(\frac{4+\log 5}{2\log 5}\right) \leqslant 1.56,$$

and $G = \log p + \log\log p + 2.15$. Because $p > 5\times 10^7$, $G > \theta$. If we let

$$U = 2e\log q\,\log(s^q)(2.15 + \log p + \log\log p)^2$$
$$= 2eq\log q(\log s)(2.15 + \log p + \log\log p)^2$$

and $C = 478$ (see [7, Sect. 6, Table 1]), we obtain from [7, Theorem 5.11] that

$$\frac{1}{p}\Lambda_2 > \exp\{-478U\};$$

i.e.,

$$\Lambda_2 > p\exp\{-956eq(\log q)(\log s)(2.15 + \log p + \log\log p)^2\}. \tag{26}$$

So, by (25),

$$p \leqslant \frac{\log 4}{\log s} + 2\frac{\log q}{\log s} + 956e(2.15 + \log p + \log\log p)^2\, q(\log q). \qquad (*)$$

Now $s \geqslant q \geqslant 5$, whence, by (17),

$$p \leqslant 2.862 + 956e(2.15 + \log p + \log\log p)^2$$
$$\times (32.91 + 3\log\log p)(e^{32.91}(\log p)^3).$$

Consequently, $p \leqslant e^{65.7} \leqslant 3.42\times 10^{28}$ and $q \leqslant e^{45.47} \leqslant 5.59\times 10^{19}$ by (17).

Finally, if $5.6\times 10^{19} \leqslant p < 10^{27}$, then we must take $f = 0.27$ and $G_0 = 2.05\log p$ in the application of [12] above. This gives $\Lambda_1 \geqslant \exp\{-e^{32.98}(\log p)^3(\log 2r)\}$. Since $p < 10^{27}$ and $r \geqslant p \geqslant 5.6\times 10^{19}$, we obtain $q \leqslant e^{33}(\log p)^3 \leqslant e^{45.4} < 5.3\times 10^{19}$.

This completes the proof of the theorem. ∎

At the other end of the spectrum, when q is small we can couple $(*)$ most fruitfully with the work of Aaltonen and Inkeri [1]. First, we rewrite $(*)$ to get

$$p \leqslant \frac{\log 4}{\log q} + 2 + 956e(2.15 + \log p + \log\log p)^2\,(q \log q). \tag{†}$$

Hence, if $q < 37$, then $p \leqslant 2 \times 10^8$.

Now, for any prime P, let h_P be the class number of the cyclotomic field $Q(\zeta_P)$, where ζ_P is a primitive Pth root of unity, and $h(-P)$ be the class number of the field $Q(\sqrt{-P})$. In [1] it is shown that if p and q are as above, then $x^p - y^q = \varepsilon$ has no solutions in the positive rational integers if p and q are odd primes for which $q^p \not\equiv q \pmod{p^2}$ and either (i) $p \nmid h_q$ or (ii) both $q \equiv 3 \pmod 4$ and $p \nmid h(-q)$. Now, by [13, p. 353], for $q < 37$, $p \nmid h_q$ if $p > q$ (since $p \nmid h_q^+$ if $p > q$ and $q < 37$). Moreover, for $5 \leqslant q < 37$ (and so $p \leqslant 2 \times 10^8$), $q^p \not\equiv q \pmod{p^2}$ except for the following pairs (see [9, p. 276], recalling that $p > q$):

q	p	q	p
5	20,771	19	43
5	40,487	19*	137
5	53,471,161	19*	63,061,489
7	491,531	23*	2,481,757
11	71	23*	13,703,077
13	863	31	79
13	1,747,591	31	6,451
17*	46,021	31*	2,806,861
17	48,947		

If we assume $p > 5 \times 10^7$, we can substitute $q = 5$ into (†). But this gives $p < 5 \times 10^7$. Therefore the case $q = 5$, $p = 53{,}471{,}161$ fails. For all pairs of entries in the table, $p^q \not\equiv p \pmod{q^2}$. In the cases with $p \equiv 3 \pmod 4$, it happens that $q \nmid h(-p)$—see [4]. Thus, by [1], only the asterisked cases remain as possiblities; i.e.,

THEOREM. *Let p, q be primes with $p > q$, and $\varepsilon = \pm 1$. Then $x^p - y^q = \varepsilon$ has no solutions in the set of positive rational integers with $q < 37$ except for the six possibilities ($q = 17$, $p = 46{,}021$), ($q = 19$, $p = 137$ or $63{,}061{,}489$), ($q = 23$, $p = 2{,}481{,}757$ or $13{,}703{,}077$), or ($q = 31$, $p = 2{,}806{,}861$).*

Although the bounds on the exponents are relatively small, the bounds on x and y are still "stratospheric." Nonetheless, the bounds on the exponents are so beguilingly low that they almost persuade one to believe that a solution to Catalan's conjecture should be possible. We hope that this short note will spur on others to complete the process.

Note Added in Proof. 1. The results in this paper were independently obtained by M. Mignotte by essentially the same proof.

2. Using the more recent work of A. Baker and G. Wustholz [*J. Reine Angew. Math.* **442** (1993), 19–62] and M. Laurent [Appendix in M. Waldschmidt's "Linear Independence of Logarithms of Algebraic Numbers," Madras Lecture Notes], in place of [12] and [7], respectively, T. Okada was able to obtain better bounds for p and q by this method (since, essentially $[Q(\sqrt{p}, \sqrt{q}, \sqrt{r/s}) : Q] = 8$).

Specifically, he shows that $p < 8.62 \times 10^{23}$ and $q < 1.18 \times 10^{17}$.

Moreover, he is also able to show that the last four pairs in the above Theorem are impossible; and that if $q \leqslant 71$, the only additional possible pairs for p, q are: $(q = 41, p = 1{,}025{,}273)$, $(q = 53, p = 97$ or $4889)$, $(q = 59, p = 2777)$, $(q = 61, p = 1861)$ and $(q = 67, p = 268{,}573)$. So there are only eight remaining possiblities for (q, p) with $q \leqslant 71$, p, q prime, $p > q$.

References

1. A. Aaltonen and K. Inkeri, Catalan's equation $x^p - y^q = 1$ and related congruences, *Math. Comp.* **56** (1991), 359–370.
2. A. Baker, Linear forms in the logarithms of algebraic numbers, *Mathematika* **13** (1966), 204–216.
3. A. Baker, Bounds for the solutions of the hyperelliptic equation, *Proc. Cambridge Philos. Soc.* **64** (1969), 439–444.
4. D. Buehl, Unpublished table of class numbers of $Q(\sqrt{-p})$.
5. J. W. S. Cassels, On the equation $a^x - b^y = 1$, II, *Proc. Cambridge Philos. Soc.* **56** (1960), 97–103; Corr., **57** (1961), 187.
6. M. Langevin, Quelques applications de nouveaux résultats de van der Poorten, *in* "Proceedings, Sém. Delange-Pisot-Poitou, 1975/1976, Paris," Exp. G12.
7. M. Mignotte and M. Waldschmidt, Linear forms in two logarithms and Schneider's method, II, *Acta Arith.* **53** (1989), 251–287.
8. T. Nagell, Des équations indéterminées $x^2 + x + 1 = y^n$ et $x^2 + x + 1 = 3y^n$, *Norsk. Mat. Forenings Skr. (1)* **2** (1921).
9. P. Ribenboim, "The Book of Prime Number Records," 2nd ed., Springer-Verlag, Heidelberg, 1989.
10. T. N. Shorey and R. Tijdeman, "Exponential Diophantine Equations," Cambridge Tracts in Mathematics, No. 87, Univ. Press, Cambridge, 1986.
11. R. Tijdeman, On the equation of Catalan, *Acta Arith.* **29** (1976), 197–209.
12. M. Waldschmidt, Minorations de combinaisons linéaires de logarithmes de nombres algébriques, *Canad. J. Math.* **45** (1993), 176–224.
13. L. C. Washington, "Introduction to Cyclotomic fields," Graduate Texts in Mathematics, Vol. 83, Springer-Verlag., Heidelberg, 1982.