REND. SEM. MAT. UNIV. PADOVA, Vol. 112 (2004)

Polynomial Squares of the Form $aX^m + b(1-X)^n + c$.

UMBERTO ZANNIER (*)

ABSTRACT - It is an open diophantine problem to prove the finiteness of the integral solutions for equations like $y^2 = 3^m + 2^n + 1$. This may be shown to correspond to one of the simplest unknown cases of a conjecture by Lang and Vojta and is probably very difficult; actually also the most basic analogous questions over function fields seem not to follow easily from known principles. With this in mind, in the present note we consider the equation $y(X)^2 = aX^m + b(1 - X)^n + c$, to be solved in the unknowns $a, b, c \in \boldsymbol{C}^*$, $m, n \in \boldsymbol{N}$ and polynomials $y \in \boldsymbol{C}[X]$; we show that only pairs (m, n) with $m, n \leqslant 8$ may give rise to solutions. Our arguments are somewhat *ad hoc* and in a way surprising; the problem is left to find a more natural and general approach.

It is a known open diophantine problem to prove the finiteness of the integral solutions to equations like

$$(*) \qquad y^2 = 3^m + 2^n + 1 \, .$$

Although certain similar equations, like e.g. $y^2 = 6^m + 2^n + 1$, can be shown to have only finitely many solutions (see [CZ] or [Z, Ch. 4]), the displayed example seems to escape from the known methods. Note that the equation represents a special, but illustrative, case of the problem of S-integral points (over $\boldsymbol{Q}$) for the variety $V := \boldsymbol{P}_2 \setminus \mathcal{D}$, where $S = \{\infty, 2, 3\}$ and where the divisor $\mathcal{D}$ is the sum of the lines $x_0 = 0$, $x_1 = 0$ and the conic $x_2^2 = x_0^2 + x_0 x_1$. The link appears after observing that an S-integral point $(1 : u : v)$ on V corresponds to an S-unit u such that $v^2 - u - 1$ is again an S-unit, namely to an equation $v^2 = 1 + u + w$ where u, w are S-units, i.e. of the form $\pm 2^a 3^b$, $a, b \in \boldsymbol{Z}$. As a special case of

(*) Indirizzo dell'A.: Dipartimento di Matematica e Informatica, Via delle Scienze, 206, 33100 Udine, Italy. E-mail: zannier@dimi.uniud.it

broad conjectures by Lang and Vojta (see [L]), one does not expect a Zariski dense set of S-integer points on such varieties $\boldsymbol{P}_2 \setminus \mathcal{D}$; the case when $\mathcal{D}$ consists of the sum of four lines is known: it leads to the so-called S-unit equation $u + v + w = 1$. The next simplest case occurs with a $\mathcal{D}$ of the mentioned shape and is still unknown (see also [Z, Ch. 4]).

J.-L. Colliot-Thélène and P. Corvaja have asked whether at least the function field analogues of the mentioned questions can be treated, i.e. replacing for instance $\boldsymbol{Q}$ with $\boldsymbol{C}(X)$ and the set S of places of $\boldsymbol{Q}$ with some finite subset of $\boldsymbol{P}_1(\boldsymbol{C})$. The simplest nontrivial choice would be $S = \{0, \infty\}$, leading to the equations $y(X)^2 = aX^m + bX^n + c$, i.e. to the easy problem of trinomials in one variable which are squares (Schinzel [Schi, 2.2.6] has described the difficult case of arbitrary k-nomials). The next case occurs with $S = \{0, 1, \infty\}$ and then the analogue of $(*)$ becomes the equation $y(X)^2 = aX^m + b(1 - X)^n + c$, for which one seeks solutions in integers m, n, nonzero complex a, b, c and polynomials $y \in \boldsymbol{C}[X]$. Note that all of this is linked to (and in fact a special case of) purely geometrical problems like the following: *Given a divisor $\mathcal{D}$ of $\boldsymbol{P}_2$, sum of two lines and a conic, classify the regular rational maps $\varphi : \boldsymbol{P}_1 \setminus \{0, 1, \infty\} \to \boldsymbol{P}_2 \setminus \mathcal{D}$.*

Now, it seems that already such basic analogue of $(*)$ escapes from the standard treatment of S-unit equations over function fields (see also Remark (a) below); hence it is tempting to investigate what sort of arguments may actually be used. After some attempts, we have found a finiteness proof by somewhat round-about methods involving roots of unity and congruences. Though the problem is certainly a special one, and though the present methods are unlikely to admit broad generalizations, for the above reasons it may be not entirely free of interest to carry out a complete proof. We formulate our conclusion as the following:

THEOREM. *There are only finitely many pairs (m, n) of positive integers such that, for some nonzero complex numbers a, b, c, the polynomial $aX^m + b(1 - X)^n + c$ is a perfect square.*

We shall give more than one argument; the last one will yield the estimate $m, n \leqslant 8$, which readily leads to a complete list of the solutions.

We thank Jean-Louis Colliot-Thélène and Pietro Corvaja for raising the problems and for several interesting conversations.

PROOF OF THEOREM. Let $a, b, c \in \boldsymbol{C}^*$ be such that $f(X) := aX^m +$

$+ b(1-X)^n + c$ is a perfect square, so all of its roots have multiplicity $\geqslant 2$. We start with the special case $m = n > 1$. Let $\xi \in \boldsymbol{C}$ be one such root, so

$$a\xi^m + b(1-\xi)^m + c = 0, \qquad a\xi^{m-1} = b(1-\xi)^{m-1}. \tag{1}$$

If ξ were a triple root, we would have also $a\xi^{m-2} + b(1-\xi)^{m-2} = 0$, which, together with the second of the equations (1), gives the false equality $\xi = \xi - 1$. Therefore all of the roots have multiplicity 2 and so there are at least $(1/2)\deg f \geqslant (m-1)/2$ distinct roots.

Combination of equations (1) yields

$$\xi^{m-1} = -c/a, \qquad (1-\xi)^{m-1} = -c/b\,.$$

Let r (resp. s) be a given complex $(m-1)$-th root of $-c/a$ (resp. $-c/b$); then $\xi = \zeta r$, $1 - \xi = \theta s$, for some $(m-1)$-th roots of unity ζ, θ, yielding the equation

$$1 - \zeta r - \theta s = 0\,.$$

But it is easy to see that for given r, $s \in \boldsymbol{C}^*$ this equation, for ζ, θ on the unit circle, represents the intersections of two distinct circles and so has at most 2 solutions in roots of unity ζ, θ. Therefore the number of possibilities for ξ is likewise bounded, leading to $m \leqslant 5$. (Note that this argument bounds by 2 the number of double roots of a polynomial $f(X)$ of the given shape, apart from the assumption that it is a perfect square.)

Let us now suppose that $m \neq n$ and that $f(X) = g(X)^2$ is the square of the polynomial $g \in \boldsymbol{C}[X]$. On putting $1 - X$ in place of X if necessary, we may assume that $m > n$, so $m = 2l$ is an even integer. Dividing $f(X)$ by $(-1)^n b$ and writing a^2 in place of $(-1)^n a/b$ and $-b$ in place of $(-1)^n c/b$, we may assume that $f(X)$ takes the shape

$$f(X) = a^2X^{2l} + (X-1)^n - b, \qquad a, b \in \boldsymbol{C}^*, \qquad 2l > n\,. \tag{2}$$

Further, by specialization we may assume that a, b are nonzero algebraic numbers and changing the sign of a if necessary we may write $g(X) = aX^l + h(X)$, where $\deg h < l$. Plugging into (2) yields

$$h(X)(2aX^l + h(X)) = (X-1)^n - b = \prod_{\zeta^n = 1} (X - 1 - \zeta u)\,,$$

where $u \in \boldsymbol{C}^*$ is a given n-th root of b. Simple inspection then leads to the

factorizations

$$2aX^l + h(X) = 2a \prod_{\zeta \in A} (X - 1 - \zeta u) \tag{3}$$

and

$$h(X) = (2a)^{-1} \prod_{\zeta \in B} (X - 1 - \zeta u), \tag{4}$$

where A, B are disjoint subsets of the set $U_n := \{\zeta \in \boldsymbol{C} : \zeta^n = 1\}$ of n-th roots of unity, such that

$$A \cup B = U_n, \qquad \#A = l, \qquad \#B = n - l .$$

In equation (3) write $1 + \theta u$ in place of X, where $\theta \in B$; then, by (4), we get $h(1 + \theta u) = 0$ and

$$(1 + \theta u)^l = u^l \prod_{\zeta \in A} (\theta - \zeta),$$

whence, dividing by u^l and setting $v = 1/u$, we find that

$$(v + \theta)^l = \prod_{\zeta \in A} (\theta - \zeta) \qquad \text{for all } \theta \in B . \tag{5}$$

In particular, v is a nonzero algebraic integer and so some conjugate of v over $\boldsymbol{Q}$ has absolute value $\geqslant 1$; but we may conjugate everything and so we may suppose from the beginning that $|v| \geqslant 1$.

An easier case now occurs when the degree $n - l$ of $h(X)$ is strictly less than $l - 1$, namely when $n < m - 1$. In this case, the coefficient of X^{l-1} on both sides of (3) vanishes, whence $\sum_{\zeta \in A} (1 + \zeta u) = 0$, i.e. $v = -(1/l) \sum_{\zeta \in A} \zeta$. The right side of this equation is plainly $\leqslant 1$ in absolute value, with equality only if $\#A = 1$. Recalling that $|v| \geqslant 1$ and that $\#A = l$, we thus conclude that $l = 1$ in this case.

Therefore we shall suppose $n = 2l - 1 = m - 1$ from now on. With this assumption we have found three essentially different ways of concluding the proof. Because of possible generalizations, we shall give all arguments, with a brief sketch of the first two and complete detail for the last one.

Sketch of first argument. The principle is to contradict (5) by an inequality, using $|v| \geqslant 1$. With this in mind, we pick $\theta_0 \in B$ so that $|v + \theta_0|$ is maximal. Then the sector of the unit circle containing $v/|v|$ and with extrema θ_0 and its symmetrical with respect to $v/|v|$, is free of elements

of B; hence the n-th roots of unity contained therein all lie in A. This information may be used to show that the right side of (5) is not too large, and must be in fact smaller than the left side, yielding a contradiction. The argument can be carried out asymptotically, taking absolute values in (5) (with $\theta = \theta_0$), then taking logarithms and finally approximating the summation over A with a suitable integral. In this last step one has first to maximize the sum in question subject to $\#A = l$ and to the mentioned constraint concerning the distribution of the set A on the unit circle. (It is easily seen that the maximum is attained when the remaining part of A lies «nearest» and symmetrically around $-\theta_0$.) The estimates one finally needs are as follows. For $0 \leqslant \omega \leqslant \pi$, introduce the function

$$G(\omega) = \pi \log\left(1 + \cos\frac{\omega}{2}\right) - \int_0^{\omega} \log(1-\cos\theta)\, d\theta - 2\int_0^{\frac{\pi-\omega}{2}} \log(1+\cos\theta)\, d\theta\,.$$

If one shows that $G(\omega)$ has a positive minimum, an upper bound for m follows (and may be effectively computed) by the skectched approach. To prove the positivity of G one calculates its first derivative, which turns out to be $H(\varrho)$, where $\varrho = \sin\frac{\omega}{2}$ and where

$$H(\varrho) = -\frac{\pi\varrho}{2(1+\sqrt{1-\varrho^2})} + \log\frac{1+\varrho}{2\varrho^2}\,.$$

One checks that: $G(0) > 0$, $G(\pi) > 0$, $G'(0^+) = +\infty$, $G'(\pi) < 0$ and $H'(\varrho) < 0$, whence $G''(\omega) < 0$. This shows that $G(\omega)$ has a unique stationary point, a local maximum, interior to the interval $[0, \pi]$. In turn, this implies that G attains its minimum either at 0 or at π; since G is positive at both points, we get what is required.

Sketch of second argument. This again exploits a similar inequality, but its deduction is more arithmetical. One first shows that if l is large enough then $v \in \boldsymbol{Q}(U_n)$. In fact, suppose the contrary and let $v' \neq v$ be a conjugate of v over $\boldsymbol{Q}(U_n)$. Conjugating (5) we see that, for an l-th root of unity $\varrho = \varrho_\theta$, we have

$$v' + \theta - \varrho v - \varrho\theta = 0\,. \tag{6}$$

We view this as a four-term linear relation among the roots of unity $1, \theta, \varrho, \varrho\theta$, with coefficients $v', 1, -v, -1$. To obtain a contradiction one applies to it a theorem of Schlickewei [S], which bounds the number

of «irreducible» solutions of linear equations in roots of unity, namely those such that no nontrivial subsum vanishes. The cases when (6) is reducible are easily dealt with as well, concluding the proof in this case. Suppose now that v is an algebraic integer in $\boldsymbol{Q}(U_n)$ and observe that for the right side of (5) we have an obvious estimate $|\prod_{\zeta\in A}(\theta-\zeta)|\leqslant c_1 c_2^l$, for every $\theta\in B$ (in fact for every θ on the unit circle), where c_1, c_2 are suitable positive absolute constants such that $c_2<2$. Then (5) implies that, for large enough l, any conjugate of $w=w_\theta:=v+\theta$ has absolute value <2. An argument of Kronecker (see [Schi], Lemma 5, p. 394) then implies that $r=\varrho+\varrho^{-1}$ for a root of unity ϱ, whence $|v+\theta|^2=(\varrho+\varrho^{-1})^2$ and

$$(|v|^2-2)+\overline{v}\theta+v\theta^{-1}-\varrho^2-\varrho^{-2}=0\,.$$

We view this as a linear relation among the five roots of unity 1, θ, θ^{-1}, ϱ^2, ϱ^{-2}, with coefficients $|v|^2-2$, $\overline{v}$, v, -1, -1. As before, the mentioned result by Schlickewei [S] implies that if l is large enough v must be a root of unity.

Finally, fix a number c_3, $c_2<c_3<2$. Then the above bound implies that if l is large enough, every conjugate w^σ of $w=w_\theta:=v+\theta$ satisfies $|w^\sigma|<c_3$. Namely, putting $\alpha:=\theta v^{-1}$ (a root of unity), we have $|1+\alpha^\sigma|<c_3$. In particular, if α has exact order N, we have $\left|1+\exp\left(\frac{2\pi i}{N}\right)\right|<c_3$. Since $c_3<2$ this plainly implies that N is bounded (in terms of c_3 only). In other words, θv^{-1} has finitely many possibilities at most, and the same then holds for θ and l, concluding the proof.

Third argument. This is completely different and exploits a congruence modulo 4. We have already noted that we may suppose that the involved coefficients are algebraic. Let $\mathcal{O}$ be a valuation ring in $\overline{\boldsymbol{Q}}$ of some place above the prime 2; factoring out we may then assume that $f(X)^2=$ $=aX^{2l}+b(X-1)^{2l-1}+c$, where $f\in\mathcal{O}[X]$ and a, b, $c\in\mathcal{O}$ do not all lie in the maximal ideal.

Note that each polynomial $P(X)$ may be expressed uniquely as a sum $A(X^2)+XB(X^2)$ of an even and an odd polynomial: in char $\neq 2$ one finds $A(X^2)=(P(X)+P(-X))/2$, $B(X^2)=(P(X)-P(-X))/2X$. Also, the denominator 2 here is in fact «apparent», since the coefficients of A and B are among the coefficients of P, so A, $B\in\mathcal{O}[X]$ if $P\in\mathcal{O}[X]$. We shall compare these expressions for the right and left sides of our equation. To

start with, one finds

$$(7) \qquad (X-1)^n = U(X^2) + XV(X^2), \qquad n = 2l-1,$$

where

$$U(X^2) = \frac{1}{2}((X-1)^n - (X+1)^n), \qquad V(X^2) = \frac{1}{2X}((X+1)^n - (1-X)^n).$$

We expand by means of the n-th roots of unity; collecting together pairs of complex conjugate terms one easily gets

$$U(X) = -\prod_{Im\zeta > 0} (2X + 2 - (X-1)(\zeta + \zeta^{-1})),$$

$$V(X) = \prod_{Im\zeta > 0} (2X + 2 + (X-1)(\zeta + \zeta^{-1})),$$

where both products are over the set S of n-th roots of unity with positive imaginary part. Note that both U, V have degree $(n-1)/2 = l-1$.

Setting $f(X) = A(X^2) + XB(X^2)$, squaring and equating even and odd parts we get

$$(8) \quad A^2(X) + XB^2(X) = aX^l + bU(X) + c, \qquad 2A(X)B(X) = bV(X).$$

Since $aX^{2l} + b(X-1)^{2l-1} + c$ is a square in $\mathcal{O}[X]$ the coefficients of odd powers of X, and in particular b, must be divisible by 2 in $\mathcal{O}$. Also, note from (7) that the leading coefficients of U and V are $-n = 1-2l$ and 1 respectively; hence they are in $\mathcal{O}^*$ and so $2 + \zeta + \zeta^{-1}$ $(Im\zeta > 0)$ and all the roots of $U(X)V(X)$ lie in $\mathcal{O}$; then the second of equations (8) easily yields (by Gauss Lemma),

$$A(X) = c_A \prod_{\zeta \in S_A} (x - \varrho_\zeta), \qquad B(X) = c_A \prod_{\zeta \in S_B} (x - \varrho_\zeta), \qquad 2c_A c_B = b,$$

where $c_A, c_B \in \mathcal{O}$, where $S_A \cup S_B = S$ is a partition and where $\varrho_\zeta = (\zeta - 2 + \zeta^{-1})/(\zeta + 2 + \zeta^{-1}) \in \mathcal{O}$. Since $\zeta + 2 + \zeta^{-1} \in \mathcal{O}^*$, this implies $\varrho_\zeta \equiv 1 \pmod 4$ for all $\zeta \in S$ and we get

$$A(X) \equiv c_A(X-1)^r, \qquad B(X) \equiv c_B(X-1)^s \pmod 4$$

where $r = \deg A$, $s = \deg B$. We have $r + s = \deg V = l - 1$; also, the first of equations (8) yields $\max(2r, 2s+1) = \deg(aX^l + bU(X) + c) = l$. Hence either $r = l/2$, $s = (l-2)/2$ or $s = (l-1)/2$ and $r = (l-1)/2 = s$.

Further, either of the formulas for $U(X)$ yields $U(X) \equiv -$

$-(2l-1)(X-1)^{l-1} \pmod 4$ and therefore the first of equations (8) implies that $aX^l + c$ is divisible by $(X-1)^h$ modulo 4, where $h = \min(2r, 2s, l-1) \geqslant l-2$, with strict inequality for odd l. If $l \geqslant 3$ this implies in particular $a + c \equiv 0 \pmod 4$, and then both a and c must be invertible in $\mathcal{O}$, because one among a, b, c is supposed to be in $\mathcal{O}^*$. Hence $X^l - 1$ is divisible by $(X-1)^{l-2}$ modulo 4, or, equivalently, $(Y+1)^l - 1$ is divisible by Y^{l-2} modulo 4. However it is easily checked that $(1+Y)^{2^q} \equiv 1 + 2Y^{2^{q-1}} + Y^{2^q} \pmod 4$ for $q \geqslant 1$, so the highest power of Y dividing $(1+Y)^l - 1$ modulo 4 is precisely 1 if l is odd and otherwise half of the highest power of 2 dividing l; since this must be $\geqslant l-2$, or even $\geqslant l-1$ for odd l, this leads to $l \leqslant 4$, which concludes the argument.

REMARKS. (a) An equation $g(X)^2 = aX^m + b(1-X)^n + c$ can be «almost» treated as a four-term S-unit equation, with methods based on differentiation, generalizing the Mason's method for the *abc*-theorem in function fields (see [BM]); however, as it stands, such an approach seems insufficient, needing just a bit of supplementary information to go through.

(b) The given proof does not allow to bound the number of double roots of a polynomial $f(X) = aX^m + b(1-X)^n + c$; namely, the assumption that $f(X)$ is a perfect square is essential for the arguments. On the one hand, this does not affect the geometrical significance of the result; on the other hand, one would expect a much stronger assertion, like e.g. an absolute bound for the number of double roots. Some evidence for this comes with the case $m = n$ when such a bound follows from our argument, as noted in the course of the proof.

(c) Things become rather easier and simpler if one assumes that a, b, c are fixed and m, n grow. Now equations (1) alone suffice to derive asymptotic relations $\xi^{m-1} \sim -c/a$, $(1-\xi)^{n-1} \sim -c/b$. In particular, all conjugates of ξ turn out to lie near some cube root of unity. Then, comparison of estimates for the discriminant leads to a contradiction for large m, n. This argument again bounds the number of double roots, and so leads to a sharper result. However assuming a, b, c to be fixed deprives the result of its geometrical implication, as in the above introductory comments.

(d) The congruence arguments, as in the third approach, may be suitably extended so to work for odd exponents n other than $2l-1$, and perhaps they might cover the whole proof. It is also possible that some variation on this method leads to a complete classifications of the regular rational maps $\varphi : \boldsymbol{P}_1 \setminus \{0, 1, \infty\} \to \boldsymbol{P}_2 \setminus \mathcal{Q}$, mentioned in the introduction.

REFERENCES

[BM] D. BROWNAWELL - D. MASSER, *Vanishing sums in function fields*, Math. Proc. Camb. Phil. Soc., **100** (1986), pp. 427-434.

[CZ] P. CORVAJA - U. ZANNIER, *On the diophantine equation* $f(a^m, y) = b^n$, Acta Arith., **94.1** (2000), pp. 25-40.

[L] S. LANG, *Number Theory III*, Encyclopoedia of Mathematical Sciences, Vol. 60, Springer-Verlag, 1991.

[Schi] A. SCHINZEL, *Polynomials with special regard to reducibility*, Cambridge Univ. Press, 2000.

[S] H. P. SCHLICKEWEI, *Equations in roots of unity*, Acta Arithmetica, **76** (1996), pp. 99-108.

[Z] U. ZANNIER, *Some Applications of Diophantine Approximations to Diophantine Equations*, Forum Editrice, Udine, 2003.

Manoscritto pervenuto in redazione il 15 aprile 2003.