Publicacions Matemàtiques, Vol **43** (1999), 207–216.

ON THE DIOPHANTINE EQUATION $x^p - x = y^q - y$

M. Mignotte[†] and A. Pethő[‡]

Abstract

We consider the diophantine equation

$$(*) \qquad x^p - x = y^q - y$$

in integers (x, p, y, q). We prove that for given p and q with $2 \leq p < q$ $(*)$ has only finitely many solutions. Assuming the abc-conjecture we can prove that p and q are bounded. In the special case $p = 2$ and y a prime power we are able to solve $(*)$ completely.

1. Introduction.

This paper was motivated by the observations of Fiedler and Alford [**FA**]. We consider the family of diophantine equations

$$(1) \qquad x^p - x = y^q - y,$$

and (except in Section 3, where we consider rational solutions) we consider only non-trivial integral solutions, that is solutions in rational integers (x, y) for which $x^p - x \neq 0$. Of course we always suppose that $|x|$, $|y|$, p, $q > 1$ and that $p \neq q$. In [**FA**], the authors give the following list of positive solutions (x, p, y, q):

$$(3, 2, 2, 3),\ (6, 2, 2, 5),\ (15, 2, 6, 3),\ (16, 2, 3, 5),$$
$$(13, 3, 3, 7),\ (91, 2, 2, 13),\ (280, 2, 5, 7),\ (4930, 2, 30, 5).$$

[†]This work was began and finished during two visits of the first author to the University of Debrecen and he wants to thank the people of this University for their kind hospitality.

[‡]Research supported in part by the Hungarian Foundation for Scientific Research, Grant No. 25157/98.

Studying systematically (1) we found no other solutions but all the above solutions, as the only solutions of some different families, except the last of the previous list. The exceptional solution $(4930, 2, 30, 5)$ corresponds to the equation $x^2 - x = y^5 - y$. It seems to be a hard problem to solve this equation.

We shall first show that non-trivial solutions can exist only when the exponents p and q are coprime. Suppose that p and q are both divisible by some prime ℓ, say $p = \ell p'$ and $q = \ell q'$, then

$$x^p - y^q = x - y = (x^{p'} - y^{q'}) \left(\sum_{i+j=\ell-1} y^{iq'} x^{jp'} \right).$$

If $\ell = 2$, we get $|x|^{p'} + |y|^{q'} \leq |x| + |y|$, contradiction. If $\ell \geq 3$, then

- if x and y are of the same sign we get again $|x|^{p'} + |y|^{q'} \leq |x| + |y|$,
- if $xy < 0$ and $y > 0$ and $x < 0$ then we see that p' must be even and we get for the third time the inequality $|x|^{p'} + |y|^{q'} \leq |x| + |y|$. The case $y < 0$ and $x > 0$ can be treated similarly. Hence we have proved

Proposition 1. *If the equation (1) has non-trivial integral solutions then p and q are coprime.*

Remark. A similar proof shows that equation (1) has no non-trivial solution (x, p, y, q) with $y = z^t$ and such that $\gcd(t, p) > 1$.

We are quite unable to prove a general result about this equation, but we prove some results about the finiteness of the set of solutions when some values among x, y, p and q are fixed. Now we study a certain collections of special cases for which we can obtain some information. Let S denote a finite set of primes. The set consisting of 1 and of all those integers which are divisible only by primes belonging to S is called the set of S-integers. Now we state the finiteness results:

0) If x and p are fixed, then as $y \mid (x^p - x)$, there are only finitely many possible solutions.
1) If $x, y \in \mathbf{Z}$ are fixed then $x^p - y^q = x - y$. This is a S-unit equation in two unknowns, hence $p,\ q < C(x, y)$, where the effectively computable function C depends only on x and y. (See the book of Shorey-Tijdeman, [**Sh-T**, Corollary 1.3].)

2) If x, y are S-integers, then (1) becomes a four-term S-unit equation. As none of x, y, $x^p - x$, $y^q - y$ and $x + y^q$ is zero, this equation has only finitely many solutions (x, p, y, q) by a theorem of Everste [**E**]. But this result is not effective.

3) If p and y are fixed, then we re-write (1) as

$$Q(x) := x^p - x + y = y^q.$$

Assume that x_0 is a multiple root of $Q(x)$. Then x_0 is also a zero of $Q'(x) = px^{p-1} - 1$, *i.e.* $x_0 = (1/p)^{\frac{1}{p-1}}\zeta$ where ζ is a $(p-1)$-th root of unity. As $Q(x_0) = 0$ we obtain

$$y = \left(\frac{1}{p}\right)^{\frac{1}{p-1}}\zeta - \left(\frac{1}{p}\right)^{\frac{p}{p-1}}\zeta = \zeta\left(\frac{1}{p}\right)^{\frac{1}{p-1}}\left(1 - \frac{1}{p}\right) = \zeta\left(\frac{1}{p}\right)^{\frac{1}{p-1}} \cdot \frac{p-1}{p}.$$

The rightmost number cannot be an integer, hence Q is square-free. As $p \geq 2$ we can apply Theorem 10.1 of [**Sh-T**] and conclude that q and x are effectively bounded.

In the next two sections we are mainly dealing with the cases p and q fixed.

2. An application of Siegel's theorem.

In this section, we use the following result of Davenport, Lewis and Schinzel [**D-L-S**]:

Theorem A. *Let $f(x)$ be a polynomial with integral coefficients of degree $n > 1$ and $g(y)$ be a polynomial with integral coefficients of degree $m > 1$. Let $D(\lambda) = \operatorname{disc}\big(f(x) + \lambda\big)$ and $E(\lambda) = \operatorname{disc}\big(g(y) + \lambda\big)$. Suppose that there are at least $\lceil n/2 \rceil$ disctinct roots of $D(\lambda) = 0$ for which $E(\lambda) \neq 0$. Then $f(x) - g(y)$ is irreducible over the complex numbers. Further, the genus of the equation $f(x) - g(y) = 0$ is strictly positive except possibly when $m = 2$ or $m = n = 3$. Apart from these possible exceptions, the equation has at most a finite number of integral solutions.*

Of course, the last assertion of this theorem is a direct application of the famous result of Siegel [**Si**] about integral points of curves of positive genus. Results which generalise the previous theorem can be found in more recent papers like [**P-S**], [**R-S**] and [**Sch**]. The book of Stepanov contains also the following more general result: Let $n = \deg(g)$, if $(n, q) = 1$ then the polynomial $y^q - y - g(x)$ is absolutely irreducible (see, [**St**, Corollary, p. 56]).

For $p = 2$ equation (1) defines an elliptic or a hyperelliptic curve according to $q = 3$ and $q > 3$. In both cases the genus of the curve is positive. Hence we may assume $2 < p < q$ in the sequel.

Put $f(x) = x^p - x$ and $g(y) = y^q - y$. First we compute the discriminant of the polynomial $h(x) = f(x) - \lambda = x^p - x - \lambda$. A common root x of $h(x)$ and $h'(x)$ satisfies $px^p - x = 0$ and $x^p - x - \lambda = 0$, hence $(p-1)x + p\lambda = 0$. This leads to the formula $D(\lambda) := \mathrm{disc}(x^p - x - \lambda) = p^p(-\lambda)^{p-1} - (p-1)^{p-1}$. Any root λ of D satisfies

$$|\lambda| = (p-1)p^{-p/(p-1)}.$$

Since the function $z \mapsto (z-1)z^{-z/(z-1)}$ is strictly increasing for $z > 1$ [proof: derivate], the discriminants $E(\lambda) = \mathrm{disc}(g(y) - \lambda)$ and $D(\lambda)$ have no common root, and the theorem above applies. We have obtained the following result:

Theorem 1. *For given p and q with $2 \le p < q$, the diophantine equation $x^p - x = y^q - y$ has only a finite number of integral solutions.*

Assuming the *abc*-conjecture we can prove much more. Let (x, p, y, q) be a non-trivial solution of (1) with $\min\{p, q\} = p$. Then, as $|x^p| \approx |y^q|$ the *abc*-conjecture implies

$$|x^p| \le (xy(x-y))^{1+\varepsilon} < (|x|^{2+p/q})^{1+\varepsilon},$$

i.e. $p = 2$ for all but finitely many pairs (p, q).

For $p = 2$ we re-write (1) as $(2x-1)^2 = 4y^q - (4y-1)$. Applying again the *abc*-conjecture we obtain $|y^q| \approx x^2 \le (x \cdot y^2)^{1+\varepsilon}$, whence $q \le 4$ with finitely many exceptions*. For the finitely many exceptional pairs (p, q) Theorem 1 implies that $\max\{|x|, |y|\}$ is bounded. Hence we have proved

Theorem 2. *Assuming the abc-conjecture equation (1) has only finitely many non-trivial solutions.*

*We thank the referee for pointing out to this argument, and to other inaccuraces in an earlier version of this paper.

3. An application of Falting's theorem.

In this section, we look at the set of rational solutions of equation (1). Recall the following proposition (*cf.* [**Fu**, Prop. 5, p. 199]):

Proposition B. *Let $\mathcal{C}$ be an irreducible plane curve with only ordinary multiple points. Let n be the degree of $\mathcal{C}$, $r_P = m_P(\mathcal{C})$. Then the genus of $\mathcal{C}$ is given by*

$$g = \frac{(n-1)(n-2)}{2} - \sum_{P \in \mathcal{C}} \frac{r_P(r_P - 1)}{2}.$$

We consider the curve $\mathcal{C}$ defined by equation (1). It follows by Theorem A that this curve is irreducible. In the present case, for the equation

$$x^p - x = y^q - y,$$

a multiple point (x, y) would be a solution of the system

$$\begin{cases} x^p - x = y^q - y, \\ px^{p-1} - 1 = 0, \\ qy^{q-1} - 1 = 0. \end{cases}$$

This implies

$$x = p^{-\frac{1}{p-1}}, \quad y = q^{-\frac{1}{q-1}}$$

and, after some computation, we get the relation

$$\frac{1}{p^{q-1}} \left(\frac{q(p-1)}{p(q-1)} \right)^{(p-1)(q-1)} = \frac{1}{q^{p-1}}.$$

Now suppose that ℓ is a prime number which divides q and, more precisely, suppose that $\ell^\beta \| q$ and $\ell^\alpha \| p - 1$. Then ℓ does not divide p by Proposition 1, and $\ell \not| \, q - 1$. Hence

$$(\beta + \alpha)(p-1)(q-1) = -\beta(p-1),$$

which is absurd. Thus we have proved that our curve $\mathcal{C}$ does not have multiple points.

As a consequence, the genus g of $\mathcal{C}$ is equal to

$$g = \frac{(q-1)(q-2)}{2} \geq 3, \quad \text{when } q \geq 4.$$

[We suppose that $q > p$, without loss of generality.] Now, by Falting's theorem [**Fa**], we know that, there are only a finite nuber of rational points on $\mathcal{C}$. Thus we have proved the following result:

Proposition 2. *The curve $\mathcal{C}$ defined by the equation*

$$x^p - x = y^q - y$$

has only a finite number of rational points (x, y) when $2 \le p < q$ and $q \ge 4$.

Remark. In the special case $p = 2$, $q = 3$, the curve $\mathcal{C}$ is an elliptic curve with rank one, thus it contains an infinite number of rational points.

4. The case $p = 2$.

We now consider the special case $p = 2$. In this special case, multiplying by 4 gives

$$X^2 = 4y^q - 4y + 1, \tag{2}$$

where $X = 2x - 1$. Put $4y - 1 = db^2$ where d is square-free and b is a positive integer. Notice that (2) has the trivial solution $(2y-1, y, 2)$. We assume that (2) has the solutions (X, y, q), with q odd and, moreover, that y is a prime power.

Put $X' = 2y - 1$. From the relation

$$X'^2 + db^2 = 4y^2,$$

if we put

$$\alpha = \frac{1 - b\sqrt{-d}}{2}, \quad \beta = \frac{X' + b\sqrt{-d}}{2} = y - \frac{1 - b\sqrt{-d}}{2},$$

we see that α and β are algebraic integers in the number field $K := \mathbf{Q}[\sqrt{-d}]$, that $\beta = -\alpha^2$ and $y = \alpha\bar{\alpha}$ (where the bar denotes complex conjugation). Clearly $\gcd(\alpha, \bar{\alpha}) = 1$.

From the relation $X^2 + db^2 = 4y^q$, if $\gamma = \frac{X + sb\sqrt{-d}}{2}$, with $s = \pm 1$, we see that γ is an algebraic integer in K and that the ideal (γ) is equal to

$$(\gamma) = \mathbf{c}^q,$$

for some ideal $\mathbf{c}$ of the field K. This relation implies $(y) = \mathbf{c}\bar{\mathbf{c}}$. Now we use (for the first time) the fact that $y = p^f$ for some prime number p. The decomposition $(y) = \alpha\bar{\alpha}$ proves that p splits in K, say $(p) = \mathbf{p}\bar{\mathbf{p}}$ with $\mathbf{p}^f = (\alpha)$ and $\mathbf{p} \ne \bar{\mathbf{p}}$. It is easy to prove that $\gcd(\mathbf{c}, \bar{\mathbf{c}}) = 1$. Choosing s suitably we may assume that $(\gamma) = \mathbf{p}^{fq} = (\alpha)^q$.

Now, $\beta = -\alpha^2$ and $\gamma = \varepsilon\alpha^q$, where ε is a unit in K. Therefore,

$$b\sqrt{-d} = \alpha^2 - \bar{\alpha}^2 = \varepsilon\alpha^q - \bar{\varepsilon}\bar{\alpha}^q. \tag{3}$$

If $d \neq 3$ [the case $d = 1$ cannot occur: $d \equiv -1 \pmod 4$], then the only roots of unity in $Q(\sqrt{-d})$ are ± 1, thus we have

$$1 = U_2 = |U_q| \tag{4}$$

with $U_n = (\alpha^n - \bar{\alpha}^n)/(\alpha - \bar{\alpha})$.

Consider the special case $d = 3$. Then the only roots of unity in K are ± 1, $\frac{\pm 1 \pm \sqrt{-3}}{2}$. Let $\varepsilon = \frac{1+\sqrt{-3}}{2}$. Then

$$\alpha^q = \gamma \cdot \frac{1-\sqrt{-3}}{2} = \frac{X + sb\sqrt{-3}}{2} \cdot \frac{1-\sqrt{-3}}{2} = \frac{X + 3sb + (sb - X)\sqrt{-3}}{4}$$

and so

$$U_q = \frac{\alpha^q - \bar{\alpha}^q}{\alpha - \bar{\alpha}} = \frac{(sb - X)\sqrt{-3}}{2} \cdot \frac{1}{-b\sqrt{-3}} = \frac{X - sb}{2b}.$$

We have $U_q \in \mathbf{Z}$ and b odd, hence b divides X. Thus b divides also y because $4y^q = X^2 + db^2$, which implies $b = 1$ and $y = 1$, which is absurd. Hence, $\varepsilon = \frac{1+\sqrt{-3}}{2}$ is not possible. One can exclude similarly the cases $\varepsilon = \frac{1-\sqrt{-3}}{2}$ and $\varepsilon = \frac{-1\pm\sqrt{-3}}{2}$. Hence (2) implies (4) in the case $d = 3$ too.

We consider the equation (4). We have $U_0 = 0$ and $U_1 = 1$. By Theorem 4 of Beukers [**B**] equation (4) has at most two solutions in q unless $y = \alpha\bar{\alpha} = 2$, 3 and 5.

In our situation we know already two solutions of (4), namely $q = 1$ and 2, hence there are no others in the general case.

Consider the exceptional cases following Beukers [**B**]:

If $y = 2$, then $q = 1$, 2, 3, 5 and 13, and they give the (already known) solutions

$$(x, p, y, q) = (3, 2, 2, 3),\ (6, 2, 2, 5),\ (91, 2, 2, 13).$$

If $y = 3$, then $q = 1$, 2 and 5, which corresponds to the (already known) solution

$$(x, p, y, q) = (16, 2, 3, 5).$$

Finally, for $y = 5$, then $q = 1$, 2 and 7, and we obtain the (already known) solution

$$(x, p, y, q) = (280, 2, 5, 7).$$

Thus, we have proved the following result:

Theorem 3. *When y is a prime power, the diophantine equation*

$$x^2 - x = y^q - y, \quad q > 2,$$

has only the following solutions

$$(x, y, q) = (3, 2, 3),\ (6, 2, 5),\ (91, 2, 13),\ (16, 3, 5),\ (280, 5, 7).$$

We notice that the equation $x^2 = 4y^q + 4y + 1$, where y is a prime power, $q \geq 1$, $q \neq 2$, has been studied in [**T-W1**] and [**T-W2**], in the second paper it is proved that the only solutions are $(x, y, q) = (5, 3, 1)$ and $(11, 3, 3)$.

In the special case $q = 3$ we obtain the elliptic diophantine equation $x^2 - x = y^3 - y$. Mordell [**M**] proved a long time ago that it has the following set of solutions $(x, y) = (0, 0)$, $(1, 0)$, $(0, \pm 1)$, $(1, \pm 1)$, $(3, 2)$, $(-2, 2)$, $(15, 6)$, $(-14, 6)$, which contains four non-trivial solutions. For $p = 2$, by Proposition 1 the case $q = 4$ leads only to trivial solutions. Hence the following result:

Proposition 3. *For $p = 2$ and $q \leq 4$ the only non-trivial integral solutions of equation (1) are*

$$(3, 2, 2, 3),\ (15, 2, 6, 3),\ (2, 2, -2, 3),\ (-14, 2, 6, 3).$$

Remark. We are also able to treat some special cases when $p = 3$ and y is fixed. Using elliptic curves and the computer algebra system SIMATH one can also prove that for $p = 3$ and $y = 3$ the only solution (x, p, y, q) of (1) is $(13, 3, 3, 7)$, which was already found in [**FA**] and that there are no solutions for $p = 3$ and $y = 2$ or $4 \leq y \leq 8$. Here, for each fixed value of y, we have two elliptic equations to consider: $Y^2 = X^3 - X + y$ when q is even (then $Y = y^{q/2}$) and $y\,Y^2 = X^3 - X + y$ (where $Y = y^{(q-1)/2}$) when q is odd.

References

[B] F. BEUKERS, The multiplicity of binary recurrences, *Compositio Math.* **40** (1980), 251–267.

[D-L-S] H. DAVENPORT, D. J. LEWIS AND A. SCHINZEL, Equations of the form $f(x) - g(y)$, *Quart. J. Math. Oxford* **12** (1961), 304–312.

[E] J. H. EVERTSE, On sums of S-units and linear recurrences, *Compositio Math.* **53** (1984), 225–244.

[Fa] G. FALTINGS, Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, *Invent. Math.* **73** (1983), 349–366.

[FA] D. C. FIELDER AND C. O. ALFORD, Observations from computer experiments on an integer equation, in "*Seventh International Conference on Fibonacci Numbers and Their Applications*," Graz, 1996 (eds.: G. E. Bergum, A. N. Philippou and A. F. Horadam), pp. 93–103.

[Fu] W. FULTON, "*Algebraic Curves, An Introduction to Algebraic Geometry*," W. A. Benjamin Inc., 1969.

[M] L. J. MORDELL, On the integer solutions of $y(y+1)=x(x+1)(x+2)$, *Pacific J. Math.* **13** (1963), 1347–1351.

[P-S] L. PANAITOPOL AND D. ŞTEFĂNESCU, On the generalized difference polynomials, *Pacific J. Math.* **143** (1990), 341–348.

[R-S] L. A. RUBEL AND A. SCHINZEL, On difference polynomials and heredity irreducible polynomials, *J. Number Theory* **12** (1980), 230–235.

[Sch] A. SCHINZEL, Reducibility of polynomials of the form $f(x)-g(y)$, *Colloq. Math.* **18** (1967), 213–218.

[Si] C. L. SIEGEL, Über einige Anwendungen diophantischer Approximationen, *Abh. Akad. Wiss. Göttingen. Math.-Phys. Kl.* **1** (1929), 70 pp; *"Collected works"*, Springer-Verlag, Berlin, 1966, pp. 209–266.

[Sh-T] T. N. SHOREY AND R. TIJDEMAN, "*Exponential Diophantine Equations*," Cambridge Univ. Press, 1986.

[St] S. STEPANOV, "*Arithmetic of Algebraic Curves*," Monographs in Contemporary Mathematics, Consultants Bureau, New York, 1994.

[T-W1] N. TZANAKIS AND J. WOLFSKILL, On the diophantine equation $y^2 = 4q^n + 4q + 1$, *J. Number Theory* **17** (1983), 144–164.

[T-W2] N. Tzanakis and J. Wolfskill, The diophantine equation $y^2 = 4q^{a/2} + 4q + 1$, with an application to coding theory, *J. Number Theory* **17** (1983), 144–164.

M. Mignotte:
Université Louis Pasteur
Mathématique
67084 Strasbourg
FRANCE

e-mail: mignotte@math.u-strasbg.fr

A. Pethő:
Lajos Kossuth University
Institute of Mathematics and Informatics
4010 Debrecen, P.O. 12
HUNGARY

e-mail: pethoe@math.klte.hu

Primera versió rebuda el 28 d'abril de 1998,
darrera versió rebuda el 27 de gener de 1999