ACTA ARITHMETICA
136.1 (2009)

On solutions of the equation $X^n + Y^n = BZ^n$ with prime $n \mid BZ$

by

Preda Mihăilescu (Göttingen)

1. Introduction. We consider the Diophantine equation

$$X^n + Y^n = BZ^n, \tag{1}$$

a ternary equation, which generalizes Fermat's equation. As shown for instance in [BGMP, BBGH], this equation is encountered as a special case in the solution process of various Diophantine equations.

Cyclotomy approaches were established by Maillet (1901) [Ma], Lubelski (1935), Dénes (1952), Győry (1966) [G], Inkeri (1980) [I] and others; see also [G] for further references. The use of Frey curves, which led to the celebrated proof of Fermat's Last Theorem by Wiles [Wi], lays at hand, and the papers [BGMP], [BBGH] apply results on modular curves and Ribet descent to (1). Typically, modular curves can be used in the case when $\mathrm{rad}(B)$ splits in a fixed set of (small) primes. Since this leads to intricate case studies, some additional general criteria which may help eliminate some of these are called for. The purpose of this paper is to prove a new criterion, in the flavor of [G] and its generalizations mentioned in [BGMP, Section 4].

The existence of a non-trivial solution for an exponent n implies that for any prime $p \mid n$ there also exist non-trivial solutions. It suffices to consider prime exponents; for the rest of this paper, n will thus be an odd prime.

Let $\mathbb{K} = \mathbb{Q}(\zeta_n)$ be the nth cyclotomic field, h_n^+ the class number of the maximal real subfield of $\mathbb{K}$, and $\mathcal{B}_k$ the kth Bernoulli number. An assumption related to the Second Case of Fermat's Last Theorem is

$$p \nmid h_p^+, \quad p^3 \nmid \mathcal{B}_{pi}, \quad i = 2, 4, \ldots, p-3. \tag{2}$$

The condition (2) was verified by computer [BCEMS] and it holds for all primes $n < 12 \cdot 10^6$.

2000 *Mathematics Subject Classification*: Primary 11D61.
Key words and phrases: generalized Fermat equation, exponential Diophantine equations.
The research was completed while the author was holding a research chair sponsored by the Volkswagen Stiftung.

Additionally, we shall require that B is such that

$$\varphi^*(B) := \varphi(\mathrm{rad}(B)) \quad \text{and} \quad (n, \varphi^*(B)) = 1. \tag{3}$$

Here $\mathrm{rad}(B)$ is the radical of B and the condition implies that B has no prime factors $t \equiv 1 \bmod n$. In particular, none of its prime factors splits completely in the nth cyclotomic field.

As in the case of Fermat's equation, the case in which $n \mid BZ$ is more difficult for cyclotomic approaches. Theorem 4.2 in [BGMP] addresses, for instance, this case. It differs from our result by replacing condition (3) with

$$\sum_{i=1}^{r} \frac{1}{f_i} \leq \frac{n-3}{2(n-1)},$$

where $\mathrm{rad}(B) = n \prod_{i=1}^{r} p_i$ and $f_i = \mathrm{ord}_n(p_i)$; additionally, $(n-1)/f_i$ should be odd. One sees that this condition is more restrictive for the orders and the number of factors of B than (3). Furthermore, if $(n-1)/f_i$ is odd then p_i is a quadratic residue and necessarily non-split. Thus our condition is more general. However, in Theorem 4.2 of [BGMP], it is possible that $v_n(BZ) = 2$, which is not possible in our approach, as we shall explain below.

We prove:

THEOREM 1. *Let $n > 3$ be a prime and $B > 1$ an integer satisfying* (3). *If* (1) *has a non-trivial integer solution with $n^3 \mid BZ$ and $(X, Y, Z) = 1$, then* (2) *is false*; *in particular*, $n > 12 \cdot 10^6$.

We remark that in the case when $n \nmid Z$ and $n \mid B$ it can be shown that $n^2 \mid B$. Indeed, assuming that (1) has the solution (X, Y, Z), then $X^n + Y^n \equiv 0 \bmod n$, so $X + Y \equiv 0 \bmod n$. Thus $X^n + Y^n = X^n + (nT - X)^n \equiv 0 \bmod n^2$. In particular, if (1) has a solution with $v_n(B) = 1$, then $n \mid Z$ and our condition applies. The case $v_n(B) = 2$ is not covered by this theorem. The stronger condition $n^3 \mid BZ$ implies $n^2 \mid (X + Y)$, which is required for the descent argument.

The plan of this paper is as follows: in the second section we give an overview of general facts and notions of cyclotomy which we use subsequently. In the third section we derive the adaptation of Kummer descent to the present case and indicate how further results on the First Case of Fermat's Last Theorem fit in our context. The proof of Theorem 1 follows.

2. Generalities. The facts described in this section have been studied in the context of the cyclotomic investigation of Fermat's equation. We give here a more general frame which will be used in our context and sketch the proofs. The reader may consult [Mi, Ri] for more detail.

FACT 1. *Let n be a positive integer, $\mathbb{K} \supset \mathbb{Q}$ an Abelian field, and $a, b \in \mathcal{O}(\mathbb{K})$ coprime. Then*

$$\left(\frac{a^n + b^n}{a+b}, (a+b)\right) \Big| (n),$$

as ideals of $\mathcal{O}(\mathbb{K})$. In particular, if n is a prime and $C = (a^n + b^n)/(a+b)$, then $e = v_n(C) \in \{0, 1\}$. Here, v_n is extended to a local field in which $\mathbb{K}$ is dense.

Proof. The proof of the first statement follows from the substitution $\gamma = a + b$ and $b = \gamma - a$. For the second, note that $v_n(C) \neq 0$ implies $\gamma \equiv 0 \bmod p$ and we may write $a = -b + \nu^k \cdot \psi$, with $\nu \in \mathbb{K}$ being a uniformizer above n and $\psi \in \mathbb{K}$ with $v_n(\psi) = 0$. This leads to the claimed fact. ∎

Using the same notations, suppose that n is a prime, ζ is a primitive nth root of unity, while $e = v_n(C)$ is defined in Fact 1. Let $\mathbb{K} \subset \mathbb{L} = \mathbb{Q}(\zeta)$ and

$$\alpha = \frac{a + b\zeta}{(1-\zeta)^e} \in \mathcal{O}(\mathbb{L}) \quad \text{and} \quad t_c(\alpha) = \frac{a + b\zeta^c}{(1-\zeta^c)^e}. \tag{4}$$

In the case when $\mathbb{K} = \mathbb{Q}$, we have $\sigma_c(\alpha) = t_c(\alpha)$. We let

$$P(\alpha) = \prod_{c=1}^{n-1} t_c(\alpha) = \frac{a^n + b^n}{n^e(a+b)},$$

a product which is the norm of α when $\mathbb{K} = \mathbb{Q}$. Observe that

$$D(c) := (\alpha, t_c(\alpha)) = (1) \quad \text{for any } c \in \{2, 3, \ldots, n-1\}. \tag{5}$$

This follows from the fact that $(1-\zeta)$ is ramified and it does not divide α, while

$$(1-\zeta)^e\alpha - t_c((1-\zeta)^e\alpha) = \zeta - \zeta^c \in D(c), \quad \text{so} \quad D(c) \,|\, (1-\zeta).$$

Suppose now that

$$P(\alpha) = \frac{a^n + b^n}{n^e(a+b)} = y^n, \quad y \in \mathcal{O}(\mathbb{K}).$$

We consider the ideals $\mathfrak{A} = (\alpha, y) \subset \mathcal{O}(\mathbb{K})$ and $t_c(\mathfrak{A}) = (t_c(\alpha), y)$ which satisfy

$$\prod_{c=1}^{n-1} (t_c(\mathfrak{A})) = y \quad \text{and} \quad \mathfrak{A}^n = (\alpha). \tag{6}$$

Both identities follow from the definition, using (5). Note that if for instance $n \nmid h(\mathbb{K})$, the class number of $\mathbb{K}$, then $\mathfrak{A}$ is principal and there is a $\beta \in \mathcal{O}(\mathbb{K})$ and a unit $\varepsilon \in \mathcal{O}(\mathbb{K})^\times$ such that

$$\alpha = \varepsilon \cdot \beta^n. \tag{7}$$

The following fact is used in the context of the Second Case of Fermat's Last Theorem (see e.g. [Wa, Chapter 9]):

FACT 2. *Let $\mathbb{K} = \mathbb{Q}[\zeta]$ be the nth cyclotomic field, $n \nmid h_n^+$, and $\mathfrak{A} \subset \mathbb{Z}[\zeta]$ an ideal with $\mathfrak{A}^n = (\alpha)$ and $\alpha \in \mathbb{Z}[\zeta]$. Suppose that $\gamma_0 = \alpha/\overline{\alpha}$ is not a unit and $\gamma_0 \equiv 1 \bmod n(1-\zeta)^2$. Then $\mathfrak{A}$ is principal and $\gamma_0 = \gamma^n$ with $\gamma \in \mathbb{Q}[\zeta]$.*

The condition (3) has the following consequence:

FACT 3. *Let $\varrho, \varpi \in \mathbb{Q}[\zeta]^+$, set*

$$\mu_a = \frac{\varrho - \zeta^a \varpi}{1 - \zeta^a}, \qquad C = \frac{\varrho^n - \varpi^n}{n(\varrho - \varpi)},$$

and suppose that $(\mu_a, \mu_b) = 1$ for $a \neq b$. If $\varrho^n - \varpi^n = \beta \cdot \gamma^n$ and none of the prime ideals $\mathfrak{r} \,|\, \beta$ are totally split, then $(\beta, \mu_a) = 1$ for all $a \in \{1, \ldots, n-1\}$. In particular, $\beta \,|\, (\varrho - \varpi)$.

Proof. Note that $\overline{\mu}_a = \mu_{-a}$. If $\mathfrak{r} \,|\, \beta$ is a prime ideal such that $\mathfrak{r} \,|\, \mu_a$ for some $a > 0$ which is not totally split, let σ fix $\mathfrak{r}$. Then

$$\sigma\mathfrak{r} = \mathfrak{r} \,|\, \sigma\mu_a.$$

But since $(\mu_a, \sigma(\mu_a)) = 1$, it follows that $\mathfrak{r} = (1)$, thus proving the claim. ∎

This fact is important in our proof and requires the use of assumption (3) for our result.

3. Proof of Theorem 1. The building block of the proof is the following proposition which generalizes the idea of Kummer descent (e.g. [Wa, Chapter 9]). When $n \,|\, Z$, the presence of the parameter B in (1) distinguishes this equation from the Second Case of Fermat's Last Theorem. The condition (3) is used in order to show that the factor B is essentially unsplit during the descent process; this is due to Fact 3.

3.1. *Kummer descent*

PROPOSITION 1. *Let n be a prime, let h_n^+ be the class number of the maximal real subfield of the nth cyclotomic extension $\mathbb{K} = \mathbb{Q}(\zeta)$, and assume that (2) is satisfied for n. Let $B \in \mathbb{Z}$ be such that (3) holds, and set $\mu = (1-\zeta)(1-\overline{\zeta})$. Suppose there are $x, y, z \in \mathcal{O}(\mathbb{K}^+)$ such that $(x, y, z) = 1$ and*

$$x^n + y^n = \eta \cdot \mu^m \cdot B_1 \cdot z^n \qquad \text{with } m \geq n,\ \eta \in \mathcal{O}(\mathbb{K})^\times, \tag{8}$$

and $B_0 = B/(B, n^M) \in \mathbb{Z}$, for some arbitrarily large M, is the n-free part of B, while B_1 is such that there is an $N \geq 1$ with $B_0 \,|\, B_1 \,|\, B_0^N$. Then x, y, z are not units.

Also, there is a further triple (x', y', z') with the same properties as (x, y, z), and additionally, $z' \,|\, z$ and the number of prime ideals dividing z' is strictly smaller than the number of divisors of z.

Proof. We use the facts from Section 2 and let $\mathbb{K} = \mathbb{Q}(\zeta)$ be the nth cyclotomic extension. Since $n \mid (x+y)$, we have $e = 1$ and may define accordingly

$$\alpha = \frac{x+\zeta y}{1-\zeta} \quad \text{and} \quad C = \frac{x^n+y^n}{n(x+y)}.$$

Note that (3) implies by Fact 3 that $(B, C) = 1$ and $B \mid (x+y)$.

The product C is by definition also coprime to n and (8) implies that $C \mid z^n$. From $\alpha = -y + (x+y)/(1-\zeta)$ we deduce $(\alpha, x+y) = 1$. One verifies that $m' = v_\mu(x+y) \geq (n+1)/2$ and $\alpha/\overline{\alpha} \equiv 1 \bmod n\mu$; it follows from Fact 2 that $\alpha/\overline{\alpha} = \gamma^n$ for a $\gamma \in \mathbb{Q}[\zeta]$.

There is some ideal $I \mid (z)$ such that $(\alpha) = I^n$ and $I \cdot \overline{I}$ is real and so, by hypothesis, it is principal. Together with the fact that $\alpha/\overline{\alpha} = \gamma^n$, this yields, after some computations which are habitual in this context,

$$t_c(\alpha) = \frac{x+\zeta^c y}{1-\zeta^c} = \eta_c \cdot \varrho_c^n, \qquad \eta_c \in \mathbb{Z}[\zeta]^\times,\ \varrho_c \in \mathbb{Z}[\zeta]^+,\ c \in P, \tag{9}$$

$$\alpha_0 := \frac{x+y}{B_1 \cdot \mu^{m-n}} = \eta_0 \cdot \varrho_0^n, \qquad \eta_0 \in \mathbb{Z}[\zeta]^\times,\ \varrho_0 \in \mathbb{Z}[\zeta]^+,\ c \in P. \tag{10}$$

The second identity uses the fact that B_1 is coprime to n, while $v_n\big(\frac{x^n+y^n}{x+y}\big) = 1$. Furthermore, since x, y are real, Fact 3 applies, showing that $\big(B_1, \frac{x^n+y^n}{x+y}\big) = 1$.

From (9), $\alpha/\overline{\alpha} = (\eta/\overline{\eta})(\varrho/\overline{\varrho})^n$, and since $\alpha/\overline{\alpha} \equiv 1 \bmod n\mu$, the unit is 1. We must have $\varrho \notin \mathbb{Z}[\zeta]^\times$, since otherwise $\alpha = \overline{\alpha}$ and this would imply $(x+y)/(\overline{x+y}) = (1-\overline{\zeta})/(1-\zeta) = -\overline{\zeta}$; this is impossible since $x, y \in \mathbb{R}$. But $\mathbf{N}(\varrho) \mid z$ and it follows that z cannot be a unit, as claimed.

Let now $\phi_c = \varrho_c \cdot \varrho_{n-c}$, so $\eta_c^2 \cdot \phi_c^n = t_c(\alpha) t_{n-c}(\alpha) = -xy + (x+y)^2/|1-\zeta^c|^2$ and

$$\eta_c^2 \phi_c^n - \eta_d^2 \phi_d^n = \varepsilon \cdot \mu^{-1}(x+y)^2 = \eta' \cdot \mu^{m'} \cdot B_1^2 \cdot \varrho_0^{2n}.$$

We note that $\psi = (\eta_c/\eta_d)^2 \equiv 1 \bmod n^2 \cdot \mu^2$. It is shown in [Wa, Theorem 9.4] that the premises (2) imply that $\psi = \xi^n$ is an nth power. Thus the previous equation can be rephrased as

$$(\xi\phi_c)^n - \phi_d^n = \varepsilon\eta_d^2 \cdot \mu^{-1}(x+y)^2 = \eta' \cdot \mu^{m'} \cdot B_1^2 \cdot \varrho_0^{2n}. \tag{11}$$

Letting $x' = \xi\phi_c$, $y' = -\phi_d$ and $z' = \varrho_0^2$, we verify the equation (8): indeed, $m' \geq 2m-2$, $B_1' = B_1^2$, the unit η' and the factor $\varrho_0^2 \mid z$ are defined implicitly by the previous equation; also $(\varrho_0^2, E) = 1$ and since $\varrho_0^2 \cdot E = z$ and E is not a unit, it follows indeed that $z' \mid z$ is divisible by fewer prime ideals than z. This completes the proof. ∎

We can now prove Theorem 1:

Proof. We assume that there is a non-trivial solution of (1) and $n^3 \mid BZ$. Since $v_n\big(\frac{X^n+Y^n}{n(X+Y)}\big) = 0$, it follows that $n^2 \mid (X+Y)$. Thus the premises of (8)

are fulfilled. We can apply Proposition 1 recursively and by this variant of Kummer descent, we conclude that (8) must have solutions for a sequence of triples (x_i, y_i, z_i), $i \geq 0$, with $z_{i+1} \mid z_i$, in which the number of prime ideals dividing z_i is strictly decreasing. However, since no z_i is a unit—a fact which follows by Proposition 1 always one recursion step after the definition of z_i—this leads to a contradiction showing that (2) is false, thus completing the proof. ∎

Acknowledgments. I am grateful to Professor Kálmán Győry for a stimulating correspondence on this topic.

References

[BBGH] M. A. Bennett, N. Bruin, K. Győry and L. Hajdu, *Powers from products of consecutive terms in arithmetic progression*, Proc. London Math. Soc. (3) 92 (2006), 273–306.

[BGMP] M. A. Bennett, K. Győry, M. Mignotte and Á. Pintér, *Binomial Thue equations and polynomial powers*, Compos. Math. 142 (2006), 1103–1121.

[BGP] M. A. Bennett, K. Győry and Á. Pintér, *On the Diophantine equation* $1^k + 2^k + \ldots + x^k = y^n$, ibid. 140 (2004), 1417–1431.

[BCEMS] J. Buhler, R. Crandall, R. Ernvall, T. Metsänkylä and A. Shokrollahi, *Irregular primes and cyclotomic invariants to 12 million*, J. Symbolic Comput. 31 (2001), 89–96.

[G] K. Győry, *Über die diophantische Gleichung* $x^p + y^p = cz^p$, Publ. Math. Debrecen 13 (1966), 301–305.

[I] K. Inkeri, *On the unsolvability of some diophantine equations of a modified Fermat type*, Mathematika 27 (1980), 179–187.

[Ma] E. Maillet, *Sur les équations indéterminées de la forme* $x^\lambda + y^\lambda = cz^\lambda$, Acta Math. 24 (1901), 247–256.

[Mi] P. Mihăilescu, *Class number conditions for the diagonal case of the equation of Nagell–Ljunggren*, in: Proceedings of the W. Schmidt Jubilaeum, to appear.

[Ri] P. Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer, 1979.

[Wa] L. Washington, *Introduction to Cyclotomic Fields*, 2nd ed., Grad. Texts in Math. 83, Springer, 1996.

[Wi] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. 141 (1995), 443–551.

Mathematisches Institut der Universität Göttingen
37073 Göttingen, Germany
E-mail: preda@uni-math.gwdg.de

Received on 31.7.2007
and in revised form on 31.7.2008 (5490)