ELEMENTARY CASES OF MIHĂILESCU THEOREM

PAOLO LEONETTI

Abstract. This article aims at solving, with elementary tools, interesting cases of the well-known Mihăilescu theorem, i.e. to determine the couples of consecutive perfect powers of integers.

1. Introduction and Notations

In theory of numbers, Mihăilescu theorem is the solution of a famous and old conjecture formulated by the French mathematician Eugène Charles Catalan in 1844 (see [3]). Although it was proved in April 2002, it appeared for the first time in *Journal für die reine und angewandte Mathematik* in 2004 [12]. His formulation is really easy: there is a unique couple of powers of natural numbers such that they differ by 1. Formally:

If x, y, p, q are integers greater than 1 and $x^p - y^q = 1$ then
$x = q = 3$ and $y = p = 2$.

Since the aim is not to present the whole proof (that relies on cyclotomic fields), we show just some cases that can be completely solved with elementary tools; the interested reader can find a complete proof in [13]. Special attention has been given to collect and show all useful cases and related techniques that a student attending math contests can be interested in, with a simple and self-contained form, according to Bourbaki principles. In particular the solutions of section 2 and 3 are historically famous: the first one is due to V.A. Lebesgue [10], the second to Euler [7] and Ko Chao [9]. The case studied in section 4 represents a generalization of a whole class of problems given in math contests, e.g. when y is (the power of) a prime. Finally, sections 5 and 6 contain original results, up to my knowledge.

As usual, we write $\mathbb{Z}$ for the ordered ring of integers, $\mathbb{N}$ for the subsemiring of $\mathbb{Z}$ of nonnegative integers, and $\mathbb{P} = \{2, 3, 5, \dots\}$ for the set of all (positive rational) primes. Given a non zero integer m and a prime p, $\upsilon_p(m)$ represents the p-adic valuation of m, i.e. the (unique) non negative integer k such that $p^k \mid m$ but $p^{k+1} \nmid m$; moreover $\text{gpf}(m)$ represents the greatest prime p such that $p \mid m$, whenever $m \geq 2$, and assume that $\text{gpf}(1) = 1$. For notation and terminology used but not defined here, as well as for material concerning classical topics in number theory, the reader should refer to [8].

Notice finally that if (x, y, p, q) is a solution to

$$x^p - y^q = 1 \text{ such that } \min\{x, y, p, q\} \geq 2 \tag{1}$$

then $\left(x^{p/\text{gpf(p/gpf(p))}}, y^{q/\text{gpf(q/gpf(q))}}, \frac{p}{\text{gpf(p/gpf(p))}}, \frac{q}{\text{gpf(q/gpf(q))}}\right)$ is a solution of (1) too. That's why from now on we can assume without loss of generality that p, q belong to $\mathbb{P}$.

2. Case q even

According to the remark at the end of previous section, we have to show that $y^2 + 1$ is not a power of an integer whenever $y \in \mathbb{N} \setminus \{0, 1\}$, i.e. for all $p \in \mathbb{P}$ the equation $y^2 + 1 = x^p$ does not have solutions in $\mathbb{N}$ except $(x, y) = (1, 0)$. The case $p = 2$ leads to the trivial solution, indeed x^2 and y^2 are squares that differ by 1; assume that $p \geq 3$ is a odd prime. If $2 \nmid y$ then $2 \mid x$; since $y + 1$ and $y - 1$ are two consecutive even integers, one of them will be divisible (at least) by 4, so that $8 \mid y^2 - 1$, implying that $4 \mid \frac{1}{2}(y^2 - 1) = \frac{1}{2}x^p - 1$. It means that $\frac{1}{2}x^p$ has to be odd, but it's not the case since

$$2 \leq p - 1 \leq p\upsilon_2(x) - 1 = \upsilon_2(\tfrac{1}{2}x^p) = 0,$$

implying that if a non-trivial solution exists, then $2 \mid \gcd(x - 1, y)$. Setting $q = 2$, and looking the equation (1) in $\mathbb{Z}[i]$ we have $(y + i)(y - i) = x^p$. Recallig that i is a unit, we have also that $\gcd(y + i, y - i) = \gcd(y + i, 2i)$ divides 2. But $\gcd(y + i, y - i)^2 \mid y^2 + 1 = x^p$, and by force $g \neq 2$ since $2 \nmid x$. It means $y + i$ and $y - i$ are coprime $p-$powers of some numbers in $\mathbb{Z}[i]$, i.e. there exist integers $a,b\in\mathbb{Z}$ such that

$$y+i=(a+bi)^p=\sum_{0\le j\le p}\binom{p}{j}a^j(bi)^{p-j}. \tag{2}$$

Taking only imaginary parts of equation (2) we obtain

$$1=\sum_{0\le j\le \frac{1}{2}(p-1)}\binom{p}{2j}a^{2j}b^{p-2j}i^{p-2j-1}. \tag{3}$$

Since the right hand side is divisible by b, we must have that $|b|=1$, and in particular $b^2=1$. Equation (3) becomes:

$$\sum_{0\le j\le \frac{1}{2}(p-1)}\binom{p}{2j}(-a^2)^j=(-1)^{\frac{p-1}{2}}b.$$

Directly from equation (1) we have $x^p=(a+i)^p(a-i)^p=(a^2+1)^p$ with $2\nmid x$. It means that $2\mid a$, so that

$$\sum_{0\le j\le \frac{1}{2}(p-1)}\binom{p}{2j}(-a^2)^j\equiv 1 \pmod 4 \qquad\text{implies}\qquad \sum_{0\le j\le \frac{1}{2}(p-1)}\binom{p}{2j}(-a^2)^j=1.$$

If $p=3$ the equation is clearly impossible; otherwise it can be rewritten as

$$\sum_{2\le j\le \frac{1}{2}(p-1)}\binom{p}{2j}(-a^2)^j=a^2\binom{p}{2}. \tag{4}$$

Taking in consideration the identity $\binom{p}{2j}=\binom{p}{2}\binom{p-2}{2j-2}(2j^2-j)^{-1}$ holds for all $j\in\{2,\dots,\frac{1}{2}(p-1)\}$, we can say that the following chain of inequality holds too:

$$\begin{aligned}
\upsilon_2\left(a^{2j}\binom{p}{2j}\right)&=2j\upsilon_2(a)+\upsilon_2\left(\binom{p}{2j}\right)\\
&=2j\upsilon_2(a)+\upsilon_2\left(\binom{p-2}{2j-2}\right)+\upsilon_2\left(\binom{p}{2}\right)-\upsilon_2(j)\\
&\ge 2j\upsilon_2(a)+\upsilon_2\left(\binom{p}{2}\right)-\upsilon_2(j)\\
&> 2\upsilon_2(a)+\upsilon_2\left(\binom{p}{2}\right)\\
&=\upsilon_2\left(a^2\binom{p}{2}\right).
\end{aligned}$$

It's enough to conclude that no solutions (x,y,p,q) exists to (1) whenever q is even.

3. Case p even

We already know from section 2 that if $q=2$ then the problem (1) has no solutions, so it's enough to look at two cases: $q=3$ and $q\ge 5$.

3.1. **Case** $q=3$.

According to the remark at the end of section 1, we have to solve the problem

$$x^2-y^3=1 \text{ such that } \min\{x,y\}\ge 2. \tag{5}$$

If $2\mid x$ then $\gcd(x+1,x-1)=1$ and $(x+1)(x-1)=y^3$, i.e. $x+1$ and $x-1$ are two coprime cubes that differ by 2: it's clear that no such integer $x\ge 2$ exists. Hence if (x,y) is a solution to problem (5) then $2\mid\gcd(x-1,y)$, i.e. there exist two positive integers m,n such that $x=2m+1$ and $y=2n$. Then the problem (5) can be rewritten as

$$\tfrac{1}{2}m(m+1)=n^3 \text{ such that } \min\{m,n\}\ge 1.$$

$m=1$ leads to the solution $(x,y)=(3,2)$. Now it's claimed that every triangular number greater than 1 is not a cube. If $2\mid m$ then $m=2k$ for some integer $k\ge 1$ and $k(2k+1)=n^3$: but $\gcd(k,2k+1)=1$ so that k and $2k+1$ need to be both cubes; otherwise $2\nmid m$, i.e. $m=2k-1$ for some integer $k\ge 2$ and $k(2k-1)=n^3$:

again, $\gcd(k, 2k-1) = 1$ so that k and $2k-1$ need to be both cubes. So, once fixed $\ell \in \{-1, 1\}$ we have a system in $\mathbb{Z}$ of the form $k = \alpha^3$ and $2k + \ell = \beta^3$ with $\gcd(\alpha, \beta) = 1$. Since $\ell = \ell^3$, then the previous system is equivalent to $(\beta\alpha^{-1})^3 + (-\ell\alpha^{-1})^3 = 2$. It means that it's *sufficient* to show that the equation $\upsilon^3 + \tau^3 = 2\mu^3$ has no solutions in $\mathbb{N}$, except the trivial ones $\upsilon = \tau = \mu$. According to [5], it has been proved that the equation

$$\lambda^3 + \tau^3 = 2^n \mu^3 \tag{6}$$

has no integer non trivial solutions for all integers $n \in \mathbb{N}$. In particular, Euler proved it for $n \in \{0, 1\}$ in [7], and Dirichlet concluded by descent the impossibility of (6) for all integers $n \geq 2$ (see [6]). Let's produce anyway a sketch of the proof for the case $n = 1$ (a detailed version can be found e.g. in [14]).

If (λ, τ, μ) is a solution of (6) then $(\lambda\gcd(\lambda, \tau)^{-1}, \tau\gcd(\lambda, \tau)^{-1}, \mu\gcd(\lambda, \tau)^{-1})$ is a solution too: that is why we can assume without loss of generality that $\gcd(\lambda, \tau) = 1$ and $\lambda \leq \tau$ since the equation is symmetric. In particular $2 \nmid \lambda\tau$, so we can define two non negative integers $u = \frac{1}{2}(\lambda + \tau)$ and $v = \frac{1}{2}(\lambda - \tau)$ such that $\gcd(u, v) = 1$ and $u(u^2 + 3v^2) = \mu^3$: we have to show if $v \geq 1$ then no solution exists.

- If $3 \nmid u$ then $\gcd(u, u^2 + 3v^2) = 1$, so there exist integers z_1, z_2 such that $\gcd(z_1, z_2) = 1$, $u = z_1^3$, and $u^2 + 3v^2 = z_2^3$; once defined the integer $t = z_2 - z_1^2$, we obtain $t(t^2 + 3tz_1^2 + 3z_2) = 3v^2$. Looking it in $\mathbb{Z}/3\mathbb{Z}$ we have recursively $t = 3t_1$, $v = 3v_1$, $t_1 = 3t_2$ for some integer t_1, v_1, t_2 so that the equation can be rewritten as $t_2(27t_2^2 + 9t_2z_1^2 + z_1^4) = v_1^2$. But it's straightforward to verify that $\gcd(t_2, 27t_2^2 + 9t_2z_1^2 + z_1^4) = 1$, so they have to be both squares. It means that we end to solve in integers an equation in the form
$$\mathcal{X}^4 + 9\mathcal{X}^2\mathcal{Y}^2 + 27\mathcal{Y}^4 = \mathcal{Z}^2. \tag{7}$$
- If $3 \mid u$ then we can make some substitutions as before, i.e. $u = 3u_1$, $\mu = 3z_1$, $u_1 = 3u_2$ for some u_1, u_2, z_1 integers and we obtain that $u_2(27u_2^2 + v^2)$ is a cube; but $\gcd(u_2, 27u_2^2 + v^2) = 1$ so there exist integers χ, δ such that $u_2 = \chi^3$, $27u_2^2 + v^2 = \delta^3$. Define the new variable $\gamma = \delta - 3\chi^2$, then $\gamma(\gamma^2 + 9\chi^2\gamma + 27\chi^4)$ is a square; again, these two factors are coprime and we end with a equation in the form (7).

It means that it's enough to show that equation (7) has no solutions in integers whenever $\mathcal{X}\mathcal{Y}\mathcal{Z} \neq 0$, and without loss of generality $\gcd(\mathcal{X}, \mathcal{Y}) = 1$ (the result is well-known, see e.g. [4]). If $2 \mid \mathcal{X}$ then $4 \mid 27\mathcal{Y}^2 - \mathcal{Z}^2$, implying that $4 \mid \mathcal{Y}^2 + \mathcal{Z}^2$, i.e. $2 \mid \mathcal{Y}$ too, contradiction. If $2 \nmid \mathcal{X}\mathcal{Y}$ then $8 \mid \mathcal{Z}^2 - 5$, that is again a contradiction. It means that if $(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$ is a solution of equation (7) then $2 \nmid \mathcal{X}$ and $2 \mid \mathcal{Y}$. Define the integer $\mathcal{Y}_1 = \frac{1}{2}\mathcal{Y}$ and notice now that $3 \nmid \mathcal{X}$, otherwise $3 \mid \mathcal{Y}$ too, looking the equation in $\mathbb{Z}/3^4\mathbb{Z}$. Substituting we can rewrite the equation as

$$27\mathcal{Y}_1^4 = \left(\tfrac{1}{2}(\mathcal{Z} + \mathcal{X}^2) + 9\mathcal{Y}_1^2\right)\left(\tfrac{1}{2}(\mathcal{Z} - \mathcal{X}^2) - 9\mathcal{Y}_1^2\right). \tag{8}$$

These factors are coprime, and positive since their sum and product are both positive. We can have only two cases: in the first one $\left(\frac{1}{2}(\mathcal{Z} + \mathcal{X}^2) + 9\mathcal{Y}_1^2\right) = 27a_1^4$ and $\left(\frac{1}{2}(\mathcal{Z} - \mathcal{X}^2) - 9\mathcal{Y}_1^2\right) = b_1^4$ for some integer a_1, b_1, that implies $3 \mid 27a_1^4 - 18\mathcal{Y}_1^2 = b_1^4 + \mathcal{X}^2$, that is impossible since -1 is not a quadratic residue in $\mathbb{Z}/3\mathbb{Z}$; in the second one, $\left(\frac{1}{2}(\mathcal{Z} + \mathcal{X}^2) + 9\mathcal{Y}_1^2\right) = a_2^4$ and $\left(\frac{1}{2}(\mathcal{Z} - \mathcal{X}^2) - 9\mathcal{Y}_1^2\right) = 27b_2^4$ for some integer a_2, b_2, implying that $a_1^4 - 18\mathcal{Y}_1^2 = 27b_1^4 + \mathcal{X}^2$, with $\mathcal{Y}_1 = a_2b_2$. If $2 \mid a_2$ then we get a contradiction in $\mathbb{Z}/8\mathbb{Z}$. But a_2 and b_2 cannot be both odd since $2 \nmid \mathcal{X}$, so that by force $2 \mid b_2$. To sum up, we can rewrite the equation (8) as

$$27b_2^4 = \left(\frac{1}{2}(a_2^2 + \mathcal{X}) - \frac{9}{2}b_2^2\right)\left(\frac{1}{2}(a_2^2 - \mathcal{X}) - \frac{9}{2}b_2^2\right) \tag{9}$$

with $a_2, b_2, \mathcal{X}$ integers such that $2 \mid \gcd(a_2 - 1, b_2, \mathcal{X} - 1)$. It's straightforward to verify that factors in equation (9) are coprime and strictly positive, implying that they are (in some order) in the form $27a_3^4$ and b_3^4 for some integers a_3, a_4. But it means that $a_3^4 + 9a_3^2b_3^2 + 27b_3^4 = a_2^2$, that is again in the form of equation (7). Notice that if $\mathcal{Z} \geq 1$ then $a_2 \leq \mathcal{Y}_1 < \mathcal{Y} < \mathcal{Z}$. It implies that, assuming that $(\mathcal{X}^\star, \mathcal{Y}^\star, \mathcal{Z}^\star)$ is a solution that minimizes $\mathcal{Z}$ with $\mathcal{Z} \geq 1$ over all possible solutions of equation (7), we obtain another solution $(\mathcal{X}^{\star\star}, \mathcal{Y}^{\star\star}, \mathcal{Z}^{\star\star})$ such that $\mathcal{Z}^{\star\star} < \mathcal{Z}^\star$. Hence the unique solution of equation (7) is $(0, 0, 0)$.

3.2. **Case** $q \geq 5$.

We have the solve the problem

$$x^2 - y^q = 1 \text{ such that } \min\{x, y\} \geq 2 \tag{10}$$

where $q \geq 5$ is a prime, according to the remark at the end of section 1. If $2 \mid x$ then $\gcd(x+1, x-1) = 1$ and $(x+1)(x-1) = y^q$, i.e. $x+1$ and $x-1$ are two coprime q-powers that differ by 2: it's clear that no such integer

$x \ge 2$ exists. Hence if $q \ge 5$ is a given prime and (x, y) is a solution to problem (10) then $2 \mid \gcd(x-1, y)$. Let's prove two lemmas that will be useful for the remaining part of the proof; in particular the first one belongs to folkore, known as "Lifting the Exponent "and typically attributed to É. Lucas [11] and R.D. Carmichael [2] (the latter having fixed an error in Lucas' original work in the 2-adic case).

Lemma 1. *For all integers $m, \eta_1, \eta_2 \in \mathbb{Z}$ and $r \in \mathbb{P}$ such that $m \ge 1$, $r \nmid \eta_1\eta_2$ and $r \mid \eta_1 - \eta_2$, the following conditions are satisfied:*

- *If $r \ge 3$, then $\upsilon_r(\eta_1^m - \eta_2^m) = \upsilon_r(\eta_1 - \eta_2) + \upsilon_r(m)$.*
- *If $r = 2$ and $2 \mid m$, then $\upsilon_2(\eta_1^m - \eta_2^m) = \upsilon_2(\eta_1 - \eta_2) + \upsilon_2(\eta_1 + \eta_2) + \upsilon_2(m) - 1$.*
- *If $\upsilon_2(\eta_1 - \eta_2) \ge 2$, then $\upsilon_2(\eta_1^m - \eta_2^m) = \upsilon_2(\eta_1 - \eta_2) + \upsilon_2(m)$.*

Proof. Let's prove it by induction on $\upsilon_r(m)$. If $\upsilon_r(m) = 0$ then in $\mathbb{Z}/r\mathbb{Z}$ we have $(\eta_1^m - \eta_2^m)/(\eta_1 - \eta_2) = \sum_{0 \le i \le m-1} \eta_1^i \eta_2^{m-1-i} = m\eta_1^{m-1} \neq 0$, so that $\upsilon_r(\eta_1^m - \eta_2^m) = \upsilon_r(\eta_1 - \eta_2)$ whenever $r \nmid m$. Suppose that $r \ge 3$ and the Lemma holds for a integer m, i.e. there exists a integer ℓ such that $\gcd(\ell, r) = 1$ and $\eta_1^m - \eta_2^m = \ell r^{\upsilon_r(\eta_1-\eta_2)+\upsilon_r(m)}$. Let's verify that it still holds for rm: there exists some integer χ such that

$$\begin{aligned}
\upsilon_r\left(\frac{\eta_1^{rm} - \eta_2^{rm}}{\eta_1^m - \eta_2^m}\right) &= \upsilon_r\left(\sum_{0 \le i \le r-1} \eta_1^{mi}\eta_2^{m(r-1-i)}\right) \\
&= \upsilon_r\left(\sum_{0 \le i \le r-1} \left(\eta_2^m + \ell r^{\upsilon_r(\eta_1-\eta_2)+\upsilon_r(m)}\right)^i \eta_2^{m(r-1-i)}\right) \\
&= \upsilon_r\left(r\eta_2^{m(r-1)} + \ell r^{\upsilon_r(\eta_1-\eta_2)+\upsilon_r(m)}\eta_2^{m(r-2)} \sum_{0 \le i \le r-1} i + \chi r^2\right) \\
&= \upsilon_r\left(r\eta_2^{m(r-1)} + \frac{r-1}{2}\,\ell r^{1+\upsilon_r(\eta_1-\eta_2)+\upsilon_r(m)}\eta_2^{m(r-2)} + \chi r^2\right) \\
&= \upsilon_r\left(r\eta_2^{m(r-1)} + r^2\left(\frac{r-1}{2}\,\ell r^{-1+\upsilon_r(\eta_1-\eta_2)+\upsilon_r(m)}\eta_2^{m(r-2)} + \chi\right)\right) \\
&= 1.
\end{aligned}$$

It proves the first part of the Lemma. Suppose now that $r = 2$; if $\upsilon_2(m) = 1$ then $\upsilon_2(\eta_1^m - \eta_2^m) = \upsilon_2(\eta_1^{m/2} - \eta_2^{m/2}) + \upsilon_2(\eta_1^{m/2} - (-\eta_2)^{m/2}) = \upsilon_2(\eta_1 - \eta_2) + \upsilon_2(\eta_1 + \eta_2)$. Suppose that the Lemma holds for a even positive integer m. Let's verify that it still holds for $2m$: there exists a odd integer ϕ such that

$$\begin{aligned}
\upsilon_2\left(\frac{\eta_1^{2m} - \eta_2^{2m}}{\eta_1^m - \eta_2^m}\right) &= \upsilon_2\left(\eta_1^m + \eta_2^m\right) = \upsilon_2\left((\eta_1^m - \eta_2^m) + 2\eta_2^m\right) \\
&= \upsilon_2\left(\phi 2^{\upsilon_2(\eta_1-\eta_2)+\upsilon_2(\eta_1+\eta_2)+\upsilon_2(m)-1} + 2\eta_2^m\right) \\
&= 1.
\end{aligned}$$

Finally, the third part of the Lemma follows by previous two points $\upsilon_2(\eta_1 - \eta_2) \ge 2$ implies by force that $\upsilon_2(\eta_1 + \eta_2) = 1$. This completes the proof. □

Lemma 2. *Fix r_1, r_2 prime numbers and η_1, η_2 distinct integers such that $\max\{r_1, r_2\} \ge 3$ and $\gcd(\eta_1, \eta_2) = 1$. If there exists a integer η_3 such that $r_1 \nmid \eta_3$ and $\eta_1^{r_1} - \eta_2^{r_1} = \eta_3^{r_2}$, then there exists a integer η_4 such that $\eta_1 - \eta_2 = \eta_4^{r_2}$.*

Proof. Since $\eta_1 - \eta_2 \mid \eta_1^m - \eta_2^m$ for all integers $m \ge 1$, then $(\eta_1 - \eta_2)\left(\frac{\eta_1^{r_1} - \eta_2^{r_1}}{\eta_1 - \eta_2}\right) = \eta_3^{r_2}$. Now we have also that $\gcd\left(\eta_1 - \eta_2, \frac{\eta_1^{r_1} - \eta_2^{r_1}}{\eta_1 - \eta_2}\right) = 1$, indeed if there exists a prime r_3 such that $r_3 \mid \eta_1 - \eta_2$ then $r_3 \mid \eta_3^{r_2}$ so that by assumption $r_3 \neq r_1$; also, by Lemma 1, we obtain

$$\upsilon_{r_3}\left(\frac{\eta_1^{r_1} - \eta_2^{r_1}}{\eta_1 - \eta_2}\right) = \upsilon_{r_3}(\eta_1^{r_1} - \eta_2^{r_1}) - \upsilon_{r_3}(\eta_1 - \eta_2) = \upsilon_{r_3}(r_1) = 0.$$

If $r_2 \geq 3$ we're done because each factor has to be a r_2-power of some integer; otherwise $r_2 = 2$ and $r_1 \geq 3$ and we are left with the case $\eta_1 - \eta_2 = -\eta_4^2$ for some non zero integer η_4. But it is not possible since $\frac{\eta_1^{r_1} - \eta_2^{r_1}}{\eta_1 - \eta_2} = -\left(\frac{\eta_3}{\eta_4}\right)^2 < 0$ implies that $\eta_1 - \eta_2$ and $\eta_1^{r_1} - \eta_2^{r_1}$ have different signs. □

Since $\gcd(x+1, x-1) = 2$, we can define integers ε, a, b such that $4 \mid x - \varepsilon$, $y = 2ab$, $x + \varepsilon = 2a^q$, $x - \varepsilon = 2^{q-1}b^q$ with $\varepsilon \in \{-1, 1\}$, $\gcd(a, 2b) = 1$ and $\min\{a, b\} \geq 1$. Since $q \geq 5$ and $x \geq 2$ we obtain

$$\left(\frac{a}{b}\right)^q = 2^{q-2}\frac{x+\varepsilon}{x-\varepsilon} \geq 8\ \frac{x-1}{x+1} \geq 2,$$

implying that $a > b$. Consider now that

$$a^{2q} - (2\varepsilon b)^q = \left(\frac{1}{2}(x+\varepsilon)\right)^2 - 2\varepsilon(x-\varepsilon) = \left(\frac{1}{2}(x-3\varepsilon)\right)^2$$

If $q \nmid \frac{1}{2}(x - 3\varepsilon)$ then $a^2 - 2\varepsilon b$ has to be a square, according to Lemma 2. But it's not possible since $a^2 \neq a^2 - 2\varepsilon b$ and $|2\varepsilon b| = 2b \leq 2(a-1)$ so that $(a-1)^2 < a^2 - 2\varepsilon b < (a+1)^2$. It means that $q \mid \frac{1}{2}(x - 3\varepsilon)$ and in particular $q \nmid x$ as far as $q \geq 5$. Rewriting equation of problem (10) as $x^2 = y^q - (-1)^q$ then there exists a integer $\zeta \geq 1$ such that $y - (-1) = \zeta^2$, again by Lemma 2; in particular ζ is a odd integer and y is not square, since by assumption $y \geq 2$. It means that $(\zeta, 1)$ and $(x, y^{\frac{1}{2}(q-1)})$ are two solutions of the Pell-equation $\mathcal{A}^2 - y\mathcal{B}^2 = 1$. Looking this equation in $\mathbb{Z}[\sqrt{y}]$, there exists a integer $m \geq 1$ such that

$$x + y^{\frac{1}{2}(q-1)}\sqrt{y} = (\zeta + \sqrt{y})^m, \tag{11}$$

since $(\zeta, 1)$ is the fundamental solution (see e.g. [1] for the theory underlying Pell-equations). Looking equation (11) in $\mathbb{Z}/y\mathbb{Z}[\sqrt{y}]$ we get $x = \zeta^m + m\zeta^{m-1}\sqrt{y}$, implying that $y \mid m\zeta^{m-1}$: notice that in particular $2 \mid \gcd(y, \zeta - 1)$ implies $2 \mid m$, i.e. $\frac{1}{2}m$ is a integer. Looking finally equation (11) in $\mathbb{Z}/\zeta\mathbb{Z}[\sqrt{y}]$ we obtain $x + y^{\frac{1}{2}(q-1)}\sqrt{y} = y^{\frac{1}{2}m}$, so that $\zeta \mid y^{\frac{1}{2}(q-1)}$. Suppose that $\zeta \geq 2$, then there exists a prime r such that $r \mid k \mid y$, and by construction $r \mid y - \zeta^2 = 1$, that is a contradiction. We proved that, once fixed a prime $q \geq 5$, if (x, y, ζ, m) is a solution of (11) then $\zeta = 1$, i.e. problem (10) has no solutions.

4. Case y divides $x - 1$

According to results from section 2 and 3, if $y \mid x - 1$ we have to solve in integers the equivalent problem

$$(1 + yz)^p - y^q = 1 \text{ such that } \min\{y, z+1\} \geq 2 \text{ and } p, q \in \mathbb{P} \setminus \{2\}. \tag{12}$$

Since in $\mathbb{Z}/z\mathbb{Z}$ we have $\sum_{0 \leq i \leq p-1} (1 + yz)^i = p$ then $\gcd\left(z, \sum_{0 \leq i \leq p-1} (1 + yz)^i\right) \mid p$.

4.1. Case $\gcd\left(z, \sum_{0 \leq i \leq p-1} (1 + yz)^i\right) = 1$.

Since the equation in problem (12) can be rewritten as

$$z \sum_{0 \leq i \leq p-1} (1 + yz)^i = y^{q-1}$$

then there exist coprime integers α, β such that $\min\{\alpha, \beta - 1\} \geq 1$, $z = \alpha^{q-1}$, $\sum_{0 \leq i \leq p-1} (1 + yz)^i = \beta^{q-1}$, and $y = \alpha\beta$. It implies that $yz = \alpha^q\beta$, so that

$$\sum_{0 \leq i \leq p-1} (1 + \alpha^q\beta)^i = \beta^{q-1}.$$

Looking this equation in $\mathbb{Z}/\beta\mathbb{Z}$ we have that $\beta \mid p$, but $\beta \geq 2$ hence by force $\beta = p$. Since $q \geq 3$, we reach the contradiction looking the same equation in $\mathbb{Z}/p^2\mathbb{Z}$:

$$0 = p^{q-1} = \sum_{0 \leq i \leq p-1} (1 + \alpha^q p)^i = p + \tfrac{1}{2}p^2(p-1)\alpha^q = p.$$

4.2. **Case $\gcd\left(z, \sum_{0\le i\le p-1}(1+yz)^i\right)=p$.** There exist positive integers u,v such that $z=pu$, $\sum_{0\le i\le p-1}(1+yz)^i = pv$, so that $y^{q-1}=p^2uv$. In particular $p\mid y$ and we can define positive integers h,k such that $h=\upsilon_p(y)$ and $k=yp^{-h}$, implying that $uv=p^{h(q-1)-2}k^{q-1}$; moreover in $\mathbb{Z}/p^2\mathbb{Z}$ we have

$$\textstyle\sum_{0\le i\le p-1}(1+yz)^i=\sum_{0\le i\le p-1}(1+kzp^h)^i=p+\frac{1}{2}p^{h+1}(p-1)kz=p,$$

so that $p\nmid v$, i.e. we can define a integer $s\ge 1$ such that $\gcd(s,v)=1$, $u=p^{h(q-1)-2}s$ and $sv=k^{q-1}$. But since they are coprime, there exist positive and coprime integers ω,δ such that $s=\omega^{q-1}$ and $v=\delta^{q-1}$, implying that

$$y^{q-1}=p^{h(q-1)}\omega^{q-1}\delta^{q-1}\quad\text{ and }\quad yz=p^{hq-1}\omega^q\delta.$$

It means that we can rewrite $\sum_{0\le i\le p-1}(1+yz)^i=pv$ as $\sum_{0\le i\le p-1}(1+p^{hq-1}\omega^q\delta)^i=p\delta^{q-1}$. Multiplying both sides by $p^{hq-1}\omega^q\delta$ we get $(1+p^{hq-1}\omega^q\delta)^p-1=p^{hq}\omega^q\delta^q$, that is equivalent to

$$\sum_{1\le j\le p}\binom{p}{j}p^{j(hq-1)}\omega^{jq}\delta^j=p^{hq}\omega^q\delta^q \tag{13}$$

If $\delta\ge 2$, there exists a prime r such that $r\mid\delta$; since p does not divide $v=\delta^{q-1}$ and $\gcd(\omega,\delta)=1$ then $r\neq p$, and $\upsilon_r(p^{hq}\omega^q\delta^q)=q\upsilon_r(\delta)$. In particular equation (13) implies that

$$\upsilon_r\left(\sum_{1\le j\le p}\binom{p}{j}p^{j(hq-1)}\omega^{jq}\delta^j\right)=q\upsilon_r(\delta),$$

that is impossible since $q\ge 3$ and $\upsilon_r\left(p^{hq}\omega^q\delta\right)<\upsilon_r\left(\binom{p}{j}p^{j(hq-1)}\omega^{jq}\delta^j\right)$ for all integers $j\in\{2,\dots,p\}$. Hence $\delta=1$ and equation (13) simplifies to

$$\sum_{2\le j\le p}\binom{p}{j}p^{j(hq-1)}\omega^{jq}=0,$$

that is clearly impossible, since it's a (non-empty) sum of positive integers.

5. Case x divides q

Notice that if (x,y,p,q) is a solution to problem (1) and $x\mid q$ then $(x,y^{q/x},p,x)$ is solution too, so that we have to solve without loss of generality the following problem

$$x^p-y^x=1 \text{ such that } \min\{x,y,p\}\ge 2 \tag{14}$$

According to results in sections 2 and 3, if (x,y,p) is a solution of problem (14) then $2\nmid xp$ and $2\mid y$. It implies that $y+1\mid y^x+1=x^p$ and in particular there exists a odd prime r such that $r\mid\gcd(x,y+1)$. Thanks to Lemma 1 we have that $\upsilon_r(x^p)=\upsilon_r(y+1)+\upsilon_r(x)$, i.e. $(p-1)\upsilon_r(x)=\upsilon_r(y+1)$, so that

$$y+1\ge r^{\upsilon_r(y+1)}\ge r^{(p-1)\upsilon_r(x)}\ge 3^{p-1}\ge 2^p+1 \text{ for all } p\ge 3$$

Going back to the equation of problem (14), the following chain of inequalities holds true too

$$x^p=y^x+1=2^{px}+1>2^{px}.$$

It implies that $x>2^x$ for some integer $x\ge 2$, which is impossible.

6. Case $\gcd(y,p)=1$ and $y\le 2^p$

According to sections 2 and 3, if (x,y,p,q) is a solution to problem (1) different from $(3,2,2,3)$ then $2\nmid pq$. Thanks to Lemma 1, if $p\ge 2$ is a odd integer and a prime r divides $x-1$ then $\upsilon_r(\Phi_p(x))=\upsilon_r(p)$, where $\Phi_p(x):=\frac{x^p-1}{x-1}$. It means that, once the equation of problem (1) is rewritten as $y^q=(x-1)\Phi_p(x)$, if a prime t divides $\gcd(x-1,\Phi_p(x))$, then t divides p too; in particular $t^2\mid y^q$, but $\gcd(y,p)=1$ by assumption so the divisibility $t\mid y$ cannot be verified. Since such a prime t cannot exist, then $x-1$ and $\Phi_p(x)$ have to be coprime, and in particular there exist positive integers z,w such that $\gcd(z,w)=1$, $x-1=z^q$, $\Phi_p(x)=w^q$ and $zw=y$. Since the inequality $w^q=\Phi_p(x)>(x-1)^{p-1}=z^{q(p-1)}$ easily holds, then $w>z^{p-1}$, that is equivalent to

$y > z^p$. If we suppose that there exists a solution such that $y \le 2^p$ then by force $z = 1$, i.e. $x = 2$. As long as $2 \nmid q$, if $\upsilon_2(y+1) = 1$ then $\upsilon_2(y^q+1) = 1$, otherwise $\upsilon_2(y+1) \ge 2$ and thanks to Lemma 1 we have

$$p = \upsilon_2(x^p) = \upsilon_2(y^q+1) = \upsilon_2(y+1) + \upsilon_2(q) = \upsilon_2(y+1),$$

implying that $y + 1 \ge 2^p$ and in particular

$$2^p = y^q + 1 \ge y \cdot y^2 + 1 \ge 2y^2 + 1 \ge (y+1)^2 \ge 2^{2p},$$

that is a contradiction.

7. Acknowledgements

The author is grateful to Salvatore Tringali (Université Pierre et Marie Curie) and Carlo Fiorito de Falco (Universitá Bocconi) for suggesting remarks that improved the readability of the article.

References

[1] Barbeau, E.J., *Pell's Equation*, Problem Books in Mathematics, Springer-Verlag, 2003.
[2] Carmichael, R.D., *On the Numerical Factors of Certain Arithmetic Forms*, Amer. Math. Monthly, **16**, 1909.
[3] Catalan, E.C., *Note extraite d'une lettre adressée à l'éditeur*, J. Reine Angew. Math. 27, 1844.
[4] Cel, J., *On decomposition of a cube into the difference of two biquadrates*, Matematyka 36, 1983.
[5] Dickson, L.E., *History of the theory of numbers: Diophantine analysis. Vol. 2*, American Mathematical Soc., 1999.
[6] Dirichlet, G.L., *Werke*, II, Anhang, 352-3.
[7] Euler, L., Commentationes Arithmeticae I, *Opera Omnia*, Series I, Vol.2, Teubner, Basel, 1915.
[8] Hardy, G.H. and E.M. Wright, *An Introduction to the Theory of Numbers,* 6th edition, revised by D.R. Heath-Brown and J.H. Silverman, Oxford University Press, 2008.
[9] Ko Chao, *On the diophantine equation* $x^2 = y^n + 1$*,* $xy \neq 0$, Scientia Sinica, 1965.
[10] Lebesgue, V.A., *Sur l'impossibilité, en nombres entiers, de l'équation* $x^m = y^2 + 1$, Nouvelles Annales de Mathematiques, 1850.
[11] Lucas, É., *Théorie des Fonctions Numériques Simplement Periodiques*, Amer. J. Math, **1**, 1878.
[12] Mihăilescu, P., *Primary Cyclotomic Units and a proof of Catalan's Conjecture*, Journal für die reine und angewandte Mathematik, 572, 2004.
[13] Schoof, R., *Catalan's Conjecture*, Springer, 2007.
[14] Sierpinki, W., *Elementary theory of numbers*, Monografie Matematyczne, 1964.

Università Bocconi, via Sarfatti 25, 20100 Milano, Italy.
E-mail address: leonetti.paolo@gmail.com