MATEMATYKA
DYSKRETNA
www.ii.uj.edu.pl/preMD/

Apoloniusz TYSZKA

Small systems of Diophantine equations which have only very large integer solutions

Preprint Nr MD 052
(otrzymany dnia 5 V 2011)

Kraków
2011

Small systems of Diophantine equations which have only very large integer solutions

Apoloniusz Tyszka

Abstract. Let $f : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ be a recursively enumerable function, $E_n = \{x_i = 1,\ x_i + x_j = x_k,\ x_i \cdot x_j = x_k : i, j, k \in \{1, \ldots, n\}\}$. We prove that there is an algorithm that computes a positive integer m for which an another algorithm accepts on the input any integer $n \geq m$ and returns a system $S \subseteq E_n$ such that S has infinitely many integer solutions and each integer tuple $(x_1, \ldots, x_n)$ that solves S satisfies $x_1 = f(n)$. For each integer $n \geq 12$ we construct a system $S \subseteq E_n$ such that S has infinitely many integer solutions and they all belong to $\mathbb{Z}^n \setminus [-2^{2^{n-1}}, 2^{2^{n-1}}]^n$.

Key words and phrases: computable upper bound for the heights of integer (rational) solutions of a Diophantine equation, Davis-Putnam-Robinson-Matiyasevich theorem, Diophantine equation with a finite number of integer (rational) solutions, recursively enumerable function.

2010 Mathematics Subject Classification: 03D25, 11D99, 11U99.

This article is a shortened version of the preprint [6]. We present a general method for constructing small systems of Diophantine equations which have only very large integer solutions. Let Φ_n denote the following statement

$$\forall x_1, \ldots, x_n \in \mathbb{Z}\ \exists y_1, \ldots, y_n \in \mathbb{Z}$$

$$\left(2^{2^{n-1}} < |x_1| \Longrightarrow (|x_1| < |y_1| \vee \ldots \vee |x_1| < |y_n|)\right) \wedge$$

$$\left(\forall i, j, k \in \{1, \ldots, n\}\ (x_i + x_j = x_k \Longrightarrow y_i + y_j = y_k)\right) \wedge \tag{1}$$

$$\forall i, j, k \in \{1, \ldots, n\}\ (x_i \cdot x_j = x_k \Longrightarrow y_i \cdot y_j = y_k) \tag{2}$$

For $n \geq 2$, the bound $2^{2^{n-1}}$ cannot be decreased because for

$$(x_1, \dots, x_n) = \left(2^{2^{n-1}}, 2^{2^{n-2}}, 2^{2^{n-3}}, \dots, 256, 16, 4, 2\right)$$

the conjunction of statements (1) and (2) guarantees that

$$(y_1, \dots, y_n) = (0, \dots, 0) \vee (y_1, \dots, y_n) = \left(2^{2^{n-1}}, 2^{2^{n-2}}, 2^{2^{n-3}}, \dots, 256, 16, 4, 2\right)$$

The statement $\forall n \Phi_n$ has powerful consequences for Diophantine equations, but is still unproven, see [5]. In particular, it implies that if a Diophantine equation has only finitely many solutions in integers (non-negative integers, rationals), then their heights are bounded from above by a computable function of the degree and the coefficients of the equation. For integer solutions, this conjectural upper bound can be computed by applying equation (3) and Lemmas 2 and 7.

The statement $\forall n \Phi_n$ is equivalent to the statement $\forall n \Psi_n$, where Ψ_n denote the statement

$$\forall x_1, \dots, x_n \in \mathbb{Z}\ \exists y_1, \dots, y_n \in \mathbb{Z}$$
$$\left(2^{2^{n-1}} < |x_1| = \max(|x_1|, \dots, |x_n|) \leq 2^{2^n} \Longrightarrow (|x_1| < |y_1| \vee \dots \vee |x_1| < |y_n|)\right) \wedge$$
$$\left(\forall i, j, k \in \{1, \dots, n\}\ (x_i + x_j = x_k \Longrightarrow y_i + y_j = y_k)\right) \wedge$$
$$\forall i, j, k \in \{1, \dots, n\}\ (x_i \cdot x_j = x_k \Longrightarrow y_i \cdot y_j = y_k)$$

In contradistinction to the statements Φ_n, each statement Ψ_n can be confirmed by a brute-force search in a finite amount of time.

The statement

$$\forall n\ \forall x_1, \dots, x_n \in \mathbb{Z}\ \exists y_1, \dots, y_n \in \mathbb{Z}$$
$$(2^{2^{n-1}} < |x_1| \Longrightarrow |x_1| < |y_1|) \wedge$$
$$(\forall i, j, k \in \{1, \dots, n\}\ (x_i + x_j = x_k \Longrightarrow y_i + y_j = y_k)) \wedge$$
$$\forall i, j, k \in \{1, \dots, n\}\ (x_i \cdot x_j = x_k \Longrightarrow y_i \cdot y_j = y_k)$$

strengthens the statement $\forall n \Phi_n$ but is false, as we will show in the Corollary.

Let

$$E_n = \{x_i = 1,\ x_i + x_j = x_k,\ x_i \cdot x_j = x_k : i, j, k \in \{1, \dots, n\}\}$$

To each system $S \subseteq E_n$ we assign the system $\widetilde{S}$ defined by

$$(S \setminus \{x_i = 1 :\ i \in \{1, \dots, n\}\}) \cup$$
$$\{x_i \cdot x_j = x_j :\ i, j \in \{1, \dots, n\} \text{ and the equation } x_i = 1 \text{ belongs to } S\}$$

In other words, in order to obtain $\widetilde{S}$ we remove from S each equation $x_i = 1$ and replace it by the following n equations:

$$\begin{array}{rcl} x_i \cdot x_1 & = & x_1 \\ & \dots & \\ x_i \cdot x_n & = & x_n \end{array}$$

Lemma 1. *For each system $S \subseteq E_n$*

$$\{(x_1, \dots, x_n) \in \mathbb{Z}^n :\ (x_1, \dots, x_n) \text{ solves } \widetilde{S}\} \quad =$$
$$\{(x_1, \dots, x_n) \in \mathbb{Z}^n :\ (x_1, \dots, x_n) \text{ solves } S\} \cup \{(0, \dots, 0)\}$$

Lemma 2. *The statement Φ_n can be equivalently stated thus: if a system $S \subseteq E_n$ has only finitely many solutions in integers $x_1, \dots, x_n$, then each such solution $(x_1, \dots, x_n)$ satisfies $|x_1|, \dots, |x_n| \leq 2^{2^{n-1}}$.*

Proof. It follows from Lemma 1. □

Nevertheless, for each integer $n \geq 12$ there exists a system $S \subseteq E_n$ which has infinitely many integer solutions and they all belong to $\mathbb{Z}^n \setminus [-2^{2^{n-1}}, 2^{2^{n-1}}]^n$. We will prove it in Theorem 1. First we need few lemmas.

Lemma 3. *If a positive integer n is odd and a pair (x, y) of positive integers solves the negative Pell equation $x^2 - dy^2 = -1$, then the pair*

$$\left(\frac{\left(x + y\sqrt{d}\right)^n + \left(x - y\sqrt{d}\right)^n}{2},\ \frac{\left(x + y\sqrt{d}\right)^n - \left(x - y\sqrt{d}\right)^n}{2\sqrt{d}} \right)$$

consists of positive integers and solves the equation $x^2 - dy^2 = -1$.

Lemma 4. *The pair $(2, 1)$ solves the equation $x^2 - 5y^2 = -1$.*

Lemma 5. *If a pair (x, y) solves the equation $x^2 - 5y^2 = -1$, then the pair $(9x + 20y,\ 4x + 9y)$ solves this equation too.*

Lemma 6. *([1, p. 141, Theorem 3.4.1]) Lemmas 4 and 5 allow us to compute all positive integer solutions to $x^2 - 5y^2 = -1$.*

Theorem 1. *For each integer $n \geq 12$ there exists a system $S \subseteq E_n$ such that S has infinitely many integer solutions and they all belong to $\mathbb{Z}^n \setminus [-2^{2^{n-1}}, 2^{2^{n-1}}]^n$.*

Proof. By Lemmas 4–6, the equation $u^2 - 5v^2 = -1$ has infinitely many solutions in positive integers and all these solutions can be simply computed. For a positive integer n, let $(u(n), v(n))$ denote the n-th solution to $u^2 - 5v^2 = -1$. We define S as

$$x_1 = 1 \qquad x_1 + x_1 = x_2 \qquad x_2 + x_2 = x_3 \qquad x_1 + x_3 = x_4$$

$$x_4 \cdot x_4 = x_5 \qquad x_5 \cdot x_5 = x_6 \qquad x_6 \cdot x_7 = x_8 \qquad x_8 \cdot x_8 = x_9$$

$$x_{10} \cdot x_{10} = x_{11} \qquad x_{11} + x_1 = x_{12} \qquad x_4 \cdot x_9 = x_{12}$$

$$x_{12} \cdot x_{12} = x_{13} \qquad x_{13} \cdot x_{13} = x_{14} \qquad \ldots \qquad x_{n-1} \cdot x_{n-1} = x_n$$

The first 11 equations of S equivalently expresses that $x_{10}^2 - 5 \cdot x_8^2 = -1$ and 625 divides x_8. The equation $x_{10}^2 - 5^9 \cdot x_7^2 = -1$ expresses the same fact. Execution of the following *MuPAD* code

```
x:=2:
y:=1:
for n from 2 to 313 do
u:=9*x+20*y:
v:=4*x+9*y:
if igcd(v,625)=625 then print(n) end_if:
x:=u:
y:=v:
end_for:
float(u^2+1);
float(2^(2^(12-1)));
```

returns only $n = 313$. Therefore, in the domain of positive integers, the minimal solution to $x_{10}^2 - 5^9 \cdot x_7^2 = -1$ is given by the pair $\left(x_{10} = u(313),\ x_7 = \dfrac{v(313)}{625}\right)$. Hence, if an integer tuple $(x_1, \ldots, x_n)$ solves S, then $|x_8| \geq v(313)$ and

$$x_{12} = x_{10}^2 + 1 \geq u(313)^2 + 1 > 2^{2^{12-1}}$$

The final inequality comes from the execution of the last two instructions of the code, as they display the numbers $1.263545677e783$ and

3.231700607e616. Applying induction, we get $x_n > 2^{2^{n-1}}$. By Lemma 3 (or by [8, p. 58, Theorem 1.3.6]), the equation $x_{10}^2 - 5^9 \cdot x_7^2 = -1$ has infinitely many integer solutions. This conclusion transfers to the system S.

□

J. C. Lagarias studied the equation $x^2 - dy^2 = -1$ for $d = 5^{2n+1}$, where $n = 0, 1, 2, 3, \ldots$. His theorem says that for these values of d, the least integer solution grows exponentially with d, see [3, Appendix A].

The next theorem generalizes Theorem 1. But first we need Lemma 7 together with introductory matter.

For a Diophantine equation $D(x_1, \ldots, x_p) = 0$, let M denote the maximum of the absolute values of its coefficients. Let $\mathcal{T}$ denote the family of all polynomials $W(x_1, \ldots, x_p) \in \mathbb{Z}[x_1, \ldots, x_p]$ whose all coefficients belong to the interval $[-M, M]$ and $\deg(W, x_i) \leq d_i = \deg(D, x_i)$ for each $i \in \{1, \ldots, p\}$. Here we consider the degrees of $W(x_1, \ldots, x_p)$ and $D(x_1, \ldots, x_p)$ with respect to the variable x_i. It is easy to check that

$$\mathrm{card}(\mathcal{T}) = (2M + 1)^{(d_1 + 1) \cdot \ldots \cdot (d_p + 1)} \tag{3}$$

To each polynomial that belongs to $\mathcal{T} \setminus \{x_1, \ldots, x_p\}$ we assign a new variable x_i with $i \in \{p + 1, \ldots, \mathrm{card}(\mathcal{T})\}$. Then, $D(x_1, \ldots, x_p) = x_q$ for some $q \in \{1, \ldots, \mathrm{card}(\mathcal{T})\}$. Let $\mathcal{H}$ denote the family of all equations of the form

$$x_i = 1,\ x_i + x_j = x_k,\ x_i \cdot x_j = x_k \ \ (i, j, k \in \{1, \ldots, \mathrm{card}(\mathcal{T})\})$$

which are polynomial identities in $\mathbb{Z}[x_1, \ldots, x_p]$. If some variable x_m is assigned to a polynomial $W(x_1, \ldots, x_p) \in \mathcal{T}$, then for each ring $\boldsymbol{K}$ extending $\mathbb{Z}$ the system $\mathcal{H}$ implies $W(x_1, \ldots, x_p) = x_m$. This observation proves the following Lemma 7.

Lemma 7. *The system $\mathcal{H} \cup \{x_q + x_q = x_q\}$ is algorithmically determinable. For each ring $\boldsymbol{K}$ extending $\mathbb{Z}$, the equation $D(x_1, \ldots, x_p) = 0$ is equivalent to the system $\mathcal{H} \cup \{x_q + x_q = x_q\} \subseteq E_{\mathrm{card}(\mathcal{T})}$. Formally, this equivalence can be written as*

$$\forall x_1 \in \boldsymbol{K} \ldots \forall x_p \in \boldsymbol{K} \Big(D(x_1, \ldots, x_p) = 0 \Longleftrightarrow \exists x_{p+1}, \ldots, x_{\mathrm{card}(\mathcal{T})} \in \boldsymbol{K}$$

$$(x_1, \ldots, x_p, x_{p+1}, \ldots, x_{\mathrm{card}(\mathcal{T})}) \text{ solves the system } \mathcal{H} \cup \{x_q + x_q = x_q\} \Big)$$

For each ring $\boldsymbol{K}$ extending $\mathbb{Z}$, the equation $D(x_1, \ldots, x_p) = 0$ has only finitely many solutions in $\boldsymbol{K}$ if and only if the system $\mathcal{H} \cup \{x_q + x_q = x_q\}$ has only finitely many solutions in $\boldsymbol{K}$.

To see how Lemma 7 works in a concrete case, let us take $D(x_1, x_2) = x_1 \cdot x_2 - 1$. Then, $p = 2$, $M = 1$, $d_1 = d_2 = 1$, $\mathrm{card}(\mathcal{T}) = (2 \cdot 1 + 1)^{(1+1)\cdot(1+1)} = 3^4 = 81$. The following *MuPAD* code

```
p:=2:
M:=1:
d_1:=1: \   p
d_2:=1: / lines
mo:=[]:
for i1 from 0 to d_1 do \   p
for i2 from 0 to d_2 do / lines
mo:=append(mo,x1^(i1)*x2^(i2)): (p variables)
end_for: \   p
end_for: / lines
T:=[x1,x2]: (p variables)
for j1 from -M to M do \
for j2 from -M to M do  \ (d_1+1) ... (d_p+1)
for j3 from -M to M do  /        lines
for j4 from -M to M do /
if (j1*mo[1]+j2*mo[2]+j3*mo[3]+j4*mo[4]<>x1) and
(j1*mo[1]+j2*mo[2]+j3*mo[3]+j4*mo[4]<>x2)
then T:=append(T,j1*mo[1]+j2*mo[2]+j3*mo[3]+j4*mo[4]) end_if:
end_for: \
end_for:  \ (d_1+1) ... (d_p+1)
end_for:  /        lines
end_for: /
print(T):
for p from 1 to nops(T) do
if T[p]=1 then print(p) end_if:
end_for:
for q from 1 to nops(T) do
if T[q]=x1*x2-1 then print(q) end_if:
end_for:
H1:=[]:
H2:=[]:
for i from 1 to nops(T) do
for j from 1 to nops(T) do
for k from 1 to nops(T) do
```

```
if T[i]+T[j]=T[k] then H1:=append(H1,[i,j,k]) end_if:
end_for:
end_for:
end_for:
print(nops(H1)):
print(H1):
for i from 1 to nops(T) do
for j from 1 to nops(T) do
for k from 1 to nops(T) do
if T[i]*T[j]=T[k] then H2:=append(H2,[i,j,k]) end_if:
end_for:
end_for:
end_for:
print(nops(H2)):
print(H2):
```

first displays the list T which enumerates the elements of $\mathcal{T}$ starting from x_1 and x_2. The code finds that $T[68] = 1$ and $T[17] = x_1 \cdot x_2 - 1$. Next, the code initializes empty lists $H1$ and $H2$. In $H1$, it stores all triplets $[i, j, k]$ with $T[i] + T[j] = T[k]$. In $H2$, it stores all triplets $[i, j, k]$ with $T[i] \cdot T[j] = T[k]$. The following system

$$\begin{cases} x_{68} &= 1 \\ x_i + x_j &= x_k \quad ([i,j,k] \in H1) \\ x_i \cdot x_j &= x_k \quad ([i,j,k] \in H2) \\ x_{17} + x_{17} &= x_{17} \end{cases}$$

consists of $1 + 2401 + 485 + 1$ equations and is equivalent to $x_1 \cdot x_2 - 1 = 0$.

The Davis-Putnam-Robinson-Matiyasevich theorem states that every recursively enumerable set $\mathcal{M} \subseteq \mathbb{N}^n$ has a Diophantine representation, that is

$$(a_1, \ldots, a_n) \in \mathcal{M} \Longleftrightarrow \exists x_1, \ldots, x_m \in \mathbb{N} \;\; W(a_1, \ldots, a_n, x_1, \ldots, x_m) = 0 \qquad (4)$$

for some polynomial W with integer coefficients, see [4] and [2]. The representation (4) is algorithmically determinable, if we know a Turing machine M such that, for all $(a_1, \ldots, a_n) \in \mathbb{N}^n$, M halts on $(a_1, \ldots, a_n)$ if and only if $(a_1, \ldots, a_n) \in \mathcal{M}$, see [4] and [2].

Theorem 2. *Let $f : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ be a recursively enumerable function. Then there is an algorithm that computes a positive integer m for which an another algorithm accepts on the input any integer $n \geq m$ and returns a system $S \subseteq E_n$ such that S has infinitely many integer solutions and each integer tuple $(x_1, \ldots, x_n)$ that solves S satisfies $x_1 = f(n)$.*

Proof. By the Davis-Putnam-Robinson-Matiyasevich theorem and Lemma 7, there is an integer $s \geq 3$ such that for each integers x_1, x_2

$$x_1 = f(x_2) \Longleftrightarrow \exists x_3, \ldots, x_s \in \mathbb{Z} \;\; \Psi(x_1, x_2, \ldots, x_s) \tag{5}$$

where the formula $\Psi(x_1, x_2, \ldots, x_s)$ is algorithmically determined as a conjunction of formulae of the form $x_i = 1$, $x_i + x_j = x_k$, $x_i \cdot x_j = x_k$ $(i, j, k \in \{1, \ldots, s\})$. Let $m = 8 + 2s$, and let $[\cdot]$ denote the integer part function. For each integer $n \geq m$,

$$n - \left[\frac{n}{2}\right] - 4 - s \geq m - \left[\frac{m}{2}\right] - 4 - s \geq m - \frac{m}{2} - 4 - s = 0$$

Let S denote the following system

$$\begin{cases} \text{all equations occurring in } \Psi(x_1, x_2, \ldots, x_s) \\ n - \left[\frac{n}{2}\right] - 4 - s \text{ equations of the form } z_i = 1 \\ \begin{array}{rcl} t_1 & = & 1 \\ t_1 + t_1 & = & t_2 \\ t_2 + t_1 & = & t_3 \\ & \ldots & \\ t_{\left[\frac{n}{2}\right]-1} + t_1 & = & t_{\left[\frac{n}{2}\right]} \\ t_{\left[\frac{n}{2}\right]} + t_{\left[\frac{n}{2}\right]} & = & w \\ w + y & = & x_2 \\ y + y & = & y \text{ (if } n \text{ is even)} \\ y & = & 1 \text{ (if } n \text{ is odd)} \\ u + u & = & v \end{array} \end{cases}$$

with n variables. By equivalence (5), the system S is consistent over $\mathbb{Z}$. The equation $u + u = v$ guarantees that S has infinitely many integer solutions. If an integer n-tuple $(x_1, x_2, \ldots, x_s, \ldots, w, y, u, v)$ solves S, then by equivalence (5),

$$x_1 = f(x_2) = f(w + y) = f\left(2 \cdot \left[\frac{n}{2}\right] + y\right) = f(n)$$

□

Corollary. *Let $f : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ be a recursively enumerable function. Then there is an algorithm that computes a positive integer m for which an another algorithm accepts on the input any integer $n \geq m$ and returns an integer tuple $(x_1, \dots, x_n)$ for which $x_1 = f(n)$ and*

(6) for each integers $y_1, \dots, y_n$ the conjunction

$$\Big(\forall i \in \{1, \dots, n\} \; (x_i = 1 \Longrightarrow y_i = 1)\Big) \wedge$$

$$\Big(\forall i, j, k \in \{1, \dots, n\} \; (x_i + x_j = x_k \Longrightarrow y_i + y_j = y_k)\Big) \wedge$$
$$\forall i, j, k \in \{1, \dots, n\} \; (x_i \cdot x_j = x_k \Longrightarrow y_i \cdot y_j = y_k)$$

implies that $x_1 = y_1$.

Proof. Let $\leq_n$ denote the order on $\mathbb{Z}^n$ which ranks the tuples $(x_1, \dots, x_n)$ first according to $\max(|x_1|, \dots, |x_n|)$ and then lexicographically. The ordered set $(\mathbb{Z}^n, \leq_n)$ is isomorphic to $(\mathbb{N}, \leq)$. To compute an integer tuple $(x_1, \dots, x_n)$, we solve the system S by performing the brute-force search in the order $\leq_n$. $\square$

If $n \geq 2$, then the tuple

$$(x_1, \dots, x_n) = \left(2^{2^{n-2}}, 2^{2^{n-3}}, \dots, 256, 16, 4, 2, 1\right)$$

has property *(6)*. Unfortunately, we do not know any explicitly given integers $x_1, \dots, x_n$ with property *(6)* and $|x_1| > 2^{2^{n-2}}$.

References

[1] T. Andreescu, D. Andrica, I. Cucurezeanu, *An introduction to Diophantine equations, A problem-based approach*, Birkhäuser, Basel, 2010.

[2] L. B. Kuijer, *Creating a diophantine description of a r.e. set and on the complexity of such a description,* MSc thesis, Faculty of Mathematics and Natural Sciences, University of Groningen, 2010, `http://scripties.fwn.eldoc.ub.rug.nl`.

[3] J. C. Lagarias, *On the computational complexity of determining the solvability or unsolvability of the equation $X^2 - DY^2 = -1$,* Trans. Amer. Math. Soc. 260 (1980), no. 2, 485–508.

[4] Yu. Matiyasevich, *Hilbert's tenth problem,* MIT Press, Cambridge, MA, 1993.

[5] A. Tyszka, *A hypothetical upper bound for the solutions of a Diophantine equation with a finite number of solutions,* `http://arxiv.org/abs/0901.2093`.

[6] A. Tyszka, *Small systems of Diophantine equations which have only very large integer solutions,* `http://arxiv.org/abs/1102.4122`.

[7] A. Tyszka, *Two conjectures on the arithmetic in $\mathbb{R}$ and $\mathbb{C}$,* MLQ Math. Log. Q. 56 (2010), no. 2, 175–184.

[8] S. Y. Yan, *Number theory for computing, 2nd ed.,* Springer, Berlin, 2002.

Apoloniusz Tyszka
University of Agriculture
Faculty of Production and Power Engineering
Balicka 116B, 30-149 Kraków, Poland
E-mail address: `rttyszka@cyf-kr.edu.pl`