A characterization of Büchi's integer sequences of length 3

Pablo Sáez
Computer Science and IT Department
Universidad del Bío Bío
Chillán, Chile
E-mail: psaezg@ubiobio.cl

Xavier Vidaux
Departamento de Matemática
Facultad de Ciencias Físicas y Matemáticas
Casilla 160 C
Universidad de Concepción
E-mail: xvidaux@udec.cl

Abstract

We give a new characterization of generalized Büchi sequences (sequences whose sequence of squares has constant second difference (a), for some fixed integer a) of length 3 over the integers and a strategy for attacking Büchi's n Squares Problem. Known characterizations of integer Büchi sequences of length 3 are actually characterizations over $\mathbb{Q}$, plus some divisibility criterions that keep integer sequences.

1 Introduction and Notation

A *Büchi sequence* over a commutative ring A with unit is a sequence of elements of A whose second difference of squares is the constant sequence (2) (e.g. $(0, 7, 10)$ is a Büchi sequence). Since the first difference of a sequence of consecutive squares, e.g. $(4, 9, 16, 25)$, is a sequence of consecutive odd numbers - in the example $(5, 7, 9)$ - the second difference of such a sequence is the constant sequence (2). A Büchi sequence (x_n) is called *trivial* if there exists $x \in A$ such that for all n we have $x_n^2 = (x + n)^2$ (e.g. $(-2, 3, 4)$ and $(-4, 3, 2)$). Note that a Büchi sequence (x_1, x_2, x_3) of integers satisfies

$$x_3^2 - 2x_2^2 + x_1^2 = 2. \tag{1.1}$$

The largest known non-trivial Büchi sequences over the integers have length 4 (infinitely many such sequences are known - see for example [H2] or [PPV]). *Büchi's Problem* over a commutative ring A with unit asks whether there exists an integer M such that no non-trivial Büchi sequence of length $\geq M$ exists in A. Büchi's Problem over the integers is open. Although this problem had been studied by Büchi himself in the early seventies

2010 *Mathematics Subject Classification*: Primary 11D09.
Key words and phrases: Büchi sequences, quadratic forms.
This research was partially supported by the second author's Chilean research project FONDECYT 1090233.

(or maybe even in the sixties), it became known to the general mathematical community only after being publicized by Lipshitz [L] in 1990. Two very interesting papers on this problem by D. Hensley [H1, H2] from the early eighties' were unfortunately never published.

Though it is a very natural problem of Arithmetic, it seems that the main motivation of Büchi resided in Mathematical Logic. Indeed, he observed that if this problem had a positive answer, then using the fact that the positive existential theory of $\mathbb{Z}$ in the language of rings is undecidable (a consequence of the negative answer to Hilbert's Tenth Problem by Matiyasevic, after works by M. Davis, H. Putnam and J. Robinson - see for example [M] or [D]), he could prove that the problem of simultaneous representation of integers by a system of diagonal quadratic forms over $\mathbb{Z}$ would be undecidable (see [PPV] for a more general discussion about this aspect of Büchi's Problem).

There are various evidences that Büchi's Problem would have a positive answer over the rational numbers (hence also over the integers). First in 1980, Hensley [H1] gave a heuristic proof using counting arguments. In 2001, P. Vojta [V] gave a proof (that works actually over any number field) that depends on a conjecture by Bombieri about the locus of rational points on projective varieties of general type over a number field, giving at the same time a geometric motivation for solving Büchi's Problem. In 2009, H. Pasten proved, following Vojta, that a strong version of Büchi's Problem would have a positive answer over any number field if Bombieri's conjecture had a positive answer for surfaces - see [Pa2].

For other results related to Büchi's Problem, we refer to [PPV] and [BB].

Consider a Büchi sequence (x_1, x_2, x_3) over $\mathbb{Q}$, i.e. a sequence satisfying Equation (1.1), and write $x_2 = x_1 + u$ and $x_3 = x_1 + v$. Equation (1.1) becomes

$$(x_1 + v)^2 - 2(x_1 + u)^2 + x_1^2 = 2$$

hence

$$2vx_1 + v^2 - 4ux_1 - 2u^2 = 2.$$

So we can write x_1, x_2 and x_3 as rational functions of the variables u and v such that for any rational numbers u and v, the sequence

$$(x_1(u, v), x_2(u, v), x_3(u, v))$$

is a Büchi sequence over $\mathbb{Q}$. Writing $x_2 = x_1 + u + v$ and $x_3 = x_1 + u + 2v$ and applying the same method as above, Hensley [H2] obtains a parametrization a bit simpler that allows him to show that the sequences (x_1, x_2, x_3) over $\mathbb{Z}$ which satisfy $0 \le x_1 < x_2 < x_3$ are characterized by the above parametrization by adding the conditions that u and v are both integers and, u is even and divides $v^2 - 1$. Note that the "missing" sequences are then obtained by taking all the symmetric sequences (x_3, x_2, x_1) and adding some minus signs randomly in front of the x_i's.

In this paper, we produce a direct characterization of *generalized* Büchi sequences of length 3 over the integers (solutions over $\mathbb{Z}$ to the equation $x_3^2 - 2x_2^2 + x_1^2 = a$, where a is any fixed integer), and propose a strategy for solving Büchi's Problem.

In order to state our theorems, we need first to introduce some notation.

Notation 1.1.
- For any integer a, we will denote by Γ_a the set of integer solutions of Equation
$$x_1^2 - 2x_2^2 + x_3^2 = a \tag{1.2}$$
and by Ω_a the set of integer solutions of Equation
$$-2x_1^2 + x_2^2 - 2x_3^2 = a. \tag{1.3}$$
We will often abuse notation by identifying elements $x = (x_1, x_2, x_3)$ of Γ_a with the corresponding column matrix and elements $x = (x_1, x_2, x_3)$ of Ω_a with the row matrix $\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}$.

- Let
$$B = \begin{pmatrix} 3 & 4 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$
Note that B has determinant 1 and J has determinant -1. Indeed we have $J^{-1} = J$ and
$$B^{-1} = \begin{pmatrix} 3 & -4 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
- Let $H =< B, J >$ be the subgroup of $\mathrm{GL}_3(\mathbb{Z})$ generated by B and J.
- Write $\mathcal{C} = \{x \in \mathbb{Z}^3 \colon |x_1| \le |x_2| \text{ or } |x_1| \ge 2|x_2|\}$.
- Let
$$\Theta_a = \begin{cases} \{(x_1, x_2, x_3) \in \Gamma_a \colon |x_2| \ge \max\{|x_1|, |x_3|\}\} & \text{if } a < 0 \\ \{x \in \Gamma_a \colon x \in \mathcal{C} \text{ and } Jx \in \mathcal{C}\} & \text{if } a \ge 0 \end{cases}$$
and note that for any $x \in \Theta_a$, also $Jx \in \Theta_a$ (the equation defining Γ_a is symmetric in x_1 and x_3). Note also that each Θ_a is a subset of Γ_a.
- Let
$$\Delta_2 = \{(2, 1, 0), (-2, 1, 0), (1, 0, 1), (-1, 0, 1), (-1, 0, -1)\}$$
and note that Δ_2 is a subset of Θ_2.
- Let $\Delta'_{-2} = \{(1, 0, 0), (-1, 0, 0)\}$ and $\Delta'_1 = \{(0, 1, 0), (0, -1, 0)\}$, and note that for each $a \in \{1, -2\}$, the set Δ'_a is a subset of Ω_a.

The following theorem, proved in Section 2, consists essentially of observations, but it contains the initial ideas for this paper. The idea of using the matrix B comes from the solution of Problem 204 in Sierpiński [S].

Theorem 1.2. *The group H acts on each Γ_a by left multiplication and it acts on each Ω_a by right multiplication (in particular, the orbit of each Θ_a is included in Γ_a). Moreover, if $M \in H$ then*

1. *the first and third columns of M belong to Γ_1 and the second column of M belongs to Γ_{-2}; and*
2. *the first and third rows of M belong to Ω_{-2} and the second row of M belongs to Ω_1.*

We want to find for each integer a a set as small as possible, finite if possible, whose orbit through the action of H is exactly the set Γ_a. Next two theorems, proved in Sections 3 and 4 respectively, tell us that the sets Θ_a are good candidates.

Theorem 1.3. *For each $a \neq 0$ the set Θ_a is finite. In particular, we have*
$$\Theta_{-2} = \{(0, \pm 1, 0)\} \quad \Theta_{-1} = \{(\pm 1, \pm 1, 0), (0, \pm 1, \pm 1)\}$$
$$\Theta_0 = \{(x_1, x_2, x_3) \in \mathbb{Z}^3 \colon |x_1| = |x_2| = |x_3|\}$$
$$\Theta_1 = \{(\pm 1, 0, 0), (0, 0, \pm 1)\} \quad \Theta_2 = \{(\pm 2, \pm 1, 0), (0, \pm 1, \pm 2), (\pm 1, 0, \pm 1)\}$$
where the $\pm$ signs are independent (so for example Θ_2 has 12 elements).

Theorem 1.4. *For each integer a the orbit of Θ_a is Γ_a.*

There is some obvious (possible) redundancy in each set Θ_a: for example, for each $x \in \Theta_a$ such that $x \neq Jx$, we could take one of x or Jx out of the set. We were not able to find an optimal subset of Θ_a for each a (in a uniform way), but when $a = 2$, it is not hard to see that the set Δ_2 defined above is actually enough to generate all the sequences in Θ_2, so that we have indeed (proved in Section 4):

Corollary 1.5. *The orbit of* Δ_2 *is* Γ_2*.*

In Section 5 we will prove a series of lemmas that will allow us to show, in particular, the two following theorems in Sections 6 and 7 respectively.

Theorem 1.6. *The group* H *has presentation* $\langle x, y \mid y^2 \rangle$*, hence it is isomorphic to the free product* $\mathbb{Z} * \mathbb{Z}_2$*.*

Theorem 1.7. *Given a* 3*-terms Büchi sequence* $x = (x_1, x_2, x_3)$ *of integers there exists a matrix* $M \in H$ *and a unique* $\delta \in \Delta_2$ *such that* $x = M\delta$*. Moreover, the matrix* M *is unique with this property if* $\delta \notin \{(1, 0, 1), (-1, 0, -1)\}$*, and it is unique up to right-multiplication by* J *otherwise.*

The existence part of Theorem 1.7 is just Corollary 1.5. The fact that Δ_2 is somewhat *optimal* comes from the unicity part. In particular, there are exactly five orbits, and we show in Section 8 that in order to know in what orbit a sequence (x_1, x_2, x_3) lies, it is enough to know the residues of x_1 and of x_3 modulo 8 (see Theorem 8.1).

In Section 9, we will describe a general strategy for trying to show that all Büchi sequences of length 5 are trivial, and another strategy, that seems to be more promising, for trying to show that all Büchi sequences of length 8 are trivial.

J. Browkin suggested to us the reference [C, Section 13.5, p. 301] as it is explained how to characterize integer solutions of isotropic ternary forms through a very specific action of a subgroup of $\mathrm{GL}_2(\mathbb{Q})$. This approach has the advantage of dealing with groups that are better known than our group H, but the action itself is much less natural than ours, and it is not clear to us which of the two approaches would give a better insight into Büchi's problem. For example, the characterization of the orbits seems harder with Cassel's approach.

2 Proof of Theorem 1.2

Choose an arbitrary $x = (x_1, x_2, x_3) \in \mathbb{Z}^3$. On the one hand, the sequence $Jx = (x_3, x_2, x_1)$ (respectively xJ) is an element of Γ_a (respectively Ω_a) if and only if $x \in \Gamma_a$ (respectively $x \in \Omega_a$), since Equations (1.2) and (1.3) are symmetric in x_1 and x_3. Moreover, we have:

$$Bx = \begin{pmatrix} 3x_1 + 4x_2 \\ 2x_1 + 3x_2 \\ x_3 \end{pmatrix} \qquad \text{and} \qquad xB = \begin{pmatrix} 3x_1 + 2x_2 & 4x_1 + 3x_2 & x_3 \end{pmatrix}$$

and we have

$$x_3^2 - 2(2x_1 + 3x_2)^2 + (3x_1 + 4x_2)^2 = x_3^2 - 2x_2^2 + x_1^2$$

and

$$-2x_3^2 + (3x_1 + 2x_2)^2 - 2(4x_1 + 3x_2)^2 = -2x_3^2 + x_2^2 - 2x_1^2.$$

Hence Bx satisfies Equation (1.2) if and only if x satisfies it, and xB satisfies Equation (1.3) if and only if x satisfies it. Since J and B are in $\mathrm{GL}(3, \mathbb{Z})$, we can conclude that H acts on Γ_a and Ω_a.

Let M be a matrix in H with columns c_1, c_2 and c_3 and with rows r_1, r_2 and r_3. Since

$$M = (c_1, c_2, c_3) = \left(M \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, M \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, M \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

the columns c_1 and c_3 are in the orbit of $\Delta_1 \subseteq \Theta_1$, hence are in Γ_1, and c_2 is in the orbit of $\Delta_{-2} \subseteq \Theta_{-2}$, hence is in Γ_{-2}. Since

$$M = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} M \\ \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} M \\ \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} M \end{pmatrix}$$

the rows r_1 and r_3 are in the orbit of Δ'_{-2}, hence are in Ω_{-2}, and r_2 is in the orbit of Δ'_1, hence is in Ω_1.

3 Proof of Theorem 1.3

We separate the cases $a \geq 0$ and $a < 0$.

CASE $a \geq 0$**.** If $x \in \Theta_a$ then $x \in \mathcal{C}$ and $Jx \in \mathcal{C}$, hence we have four Cases:

1. $|x_1| \leq |x_2|$ and $|x_3| \leq |x_2|$
2. $|x_1| \leq |x_2|$ and $|x_3| \geq 2|x_2|$
3. $|x_1| \geq 2|x_2|$ and $|x_3| \leq |x_2|$
4. $|x_1| \geq 2|x_2|$ and $|x_3| \geq 2|x_2|$

Case 1: We have $x_1^2 - 2x_2^2 + x_3^2 \leq 0$, so that Equation (1.2) has no solution at all in this case, unless $a = 0$. If $a = 0$ and, either $|x_1| \neq |x_2|$ or $|x_3| \neq |x_2|$, then $0 = x_1^2 - 2x_2^2 + x_3^2 < 0$, which is absurd. Hence if $a = 0$ then $|x_1| = |x_2| = |x_3|$.

Cases 3 and 4: If $|x_1| \geq 2|x_2|$ then $2x_2^2 + x_3^2 \leq x_1^2 - 2x_2^2 + x_3^2 \leq a$ and there are only finitely many sequences that satisfy

$$2x_2^2 + x_3^2 \leq a. \tag{3.1}$$

If $a = 0$ then the only solution is $x_1 = x_2 = x_3 = 0$. Let us now find the exact solutions when $a = 1$ or $a = 2$.

Subcase (i): $|x_3| \leq |x_2|$. Equation (3.1) gives then $3x_3^2 \leq a$, hence $x_3 = 0$ and Equation (3.1) becomes $2x_2^2 \leq a$. So in the case that $a = 1$, we find $x_2 = 0$ and $1 = x_1^2 - 2x_2^2 + x_3^2 = x_1^2$, which give the solutions $(\pm 1, 0, 0)$. In the case that $a = 2$, we find that x_2^2 can be 0 or 1, but since $x_1^2 - 2x_2^2 = 2$ (by definition of Γ_2) we deduce that $x_2^2 = 1$, hence $x_1^2 = 4$, which gives the solutions $(\pm 2, \pm 1, 0)$.

Subcase (ii): $|x_3| \geq 2|x_2|$. Equation (3.1) gives then $6x_2^2 \leq a$, hence $x_2 = 0$ and Equation (3.1) becomes $x_3^2 \leq a$, so $x_3^2 \leq 1$. In the case that $a = 1$, we have $1 = x_1^2 + x_3^2$ (by definition of Γ_1) hence the solutions are of the form $(\pm 1, 0, 0)$ or $(0, 0, \pm 1)$. In the case that $a = 2$, we have $2 = x_1^2 + x_3^2$, hence $x_1^2 = x_3^2 = 1$ and the solutions are $(\pm 1, 0, \pm 1)$.

Case 2: Since the equation defining Γ_a is symmetric in x_1^2 and x_3^2, we deduce from the study of Case 3 that there are only finitely many sequences and that if $a = 0$ then the only solution is $(0, 0, 0)$. Again by symmetry, the study of Subcase (ii) of Case 3 tells us that if $a = 1$ then the solutions are of the form $(0, 0, \pm 1)$, and if $a = 2$ then the solutions are of the form $(0, \pm 1, \pm 2)$.

CASE $a < 0$**.** In this case, we have $|x_2| \geq \max\{|x_1|, |x_3|\}$, hence $x_2^2 - x_1^2$ and $x_2^2 - x_3^2$ are non-negative integers. Since

$$0 < -a = -x_1^2 + 2x_2^2 - x_3^2 = (x_2^2 - x_1^2) + (x_2^2 - x_3^2)$$

we conclude that there are only finitely many choices for $x_2^2 - x_1^2$ and $x_2^2 - x_3^2$. For each such choice, there are only finitely many choices for each x_i since $x_2^2 - x_i^2 = (x_2 - x_i)(x_2 + x_i)$. In the case that $a = -1$, we have $x_2^2 - x_1^2 = 1$ and $x_2^2 - x_3^2 = 0$, which gives the solutions $(0, \pm 1, \pm 1)$, or the symmetric case that gives the solutions $(0, \pm 1, \pm 1)$. Assume now that $a = -2$. Since a difference of two squares cannot be 2, we have $x_2^2 - x_1^2 = 1$ and $x_2^2 - x_3^2 = 1$, which gives the solutions $(0, \pm 1, 0)$.

4 Proof of Theorem 1.4

The idea is to define a function $\varphi\colon \Gamma_a \to \Gamma_a$ constant on Θ_a, involving only J, B and B^{-1}, such that, given $x = (x_1, x_2, x_3) \in \Gamma_a$ there exists a positive integer n depending only on x such that the n-th iterate $\varphi^n(x)$ belongs to Θ_a (where φ^n denotes the function φ composed n times with itself).

Recalling that $\mathcal{C} = \{x \in \mathbb{Z}^3 \colon |x_1| \le |x_2| \text{ or } |x_1| \ge 2|x_2|\}$, the following four sets

$$\begin{aligned}&\Theta_a\\ \Gamma_a^1 &= \{x \in \Gamma_a \setminus \Theta_a \colon x \notin \mathcal{C} \text{ and } x_1x_2 > 0\}\\ \Gamma_a^0 &= \{x \in \Gamma_a \setminus \Theta_a \colon x \in \mathcal{C}\}\\ \Gamma_a^{-1} &= \{x \in \Gamma_a \setminus \Theta_a \colon x \notin \mathcal{C} \text{ and } x_1x_2 < 0\}\end{aligned}$$

form a partition of Γ_a (observe that if x_1 or x_2 is 0 then x is in Γ_a^0).

The function φ_a is defined in the following way:

$$\varphi_a(x) = \begin{cases} x & \text{if } x \in \Theta_a \\ Jx & \text{if } x \in \Gamma_a^0 \\ B^{-1}x & \text{if } x \in \Gamma_a^1 \\ Bx & \text{if } x \in \Gamma_a^{-1}. \end{cases}$$

Notation 4.1. If $x = (x_1, x_2, x_3) \in \Gamma_a$ then we will write

$$\varphi_a(x) = (\varphi_a(x)_1, \varphi_a(x)_2, \varphi_a(x)_3).$$

The following lemma finishes the proof of the Theorem.

Lemma 4.2. *Let $x = (x_1, x_2, x_3) \in \Gamma_a$. We have:*

1. *Θ_a is fixed by φ_a;*
2. *$\varphi_a(\Gamma_a^0) \subseteq \Theta_a \cup \Gamma_a^1 \cup \Gamma_a^{-1}$, and if $x \in \Gamma_a^0$ then $\varphi_a(x)_2 = x_2$; and*
3. *if $x \in \Gamma_a^1 \cup \Gamma_a^{-1}$ then $|\varphi_a(x)_2| < |x_2|$.*

Therefore, for all $x \in \Gamma_a$ there exists a positive integer n such that: for all integer $m \ge n$ we have $\varphi_a^m(x) = \varphi_a^n(x) \in \Theta_a$.

Proof. Assume that the three items have been proven and let $x \in \Gamma_a \setminus \Theta_a$. Applying Items 2 and 3 of the lemma repeatedly, the second term of the sequence decreases in absolute value until getting to an element of Θ_a, and the conclusion of the lemma follows.

Let us now prove each item.

1. By definition of φ_a.
2. If $x = (x_1, x_2, x_3) \in \Gamma_a^0$ then $\varphi_a(x) = Jx = (x_3, x_2, x_1)$, hence trivially $\varphi_a(x)_2 = x_2$. In order to obtain a contradiction, suppose $\varphi_a(x) \in \Gamma_a^0$, so that we have: $x \in \mathcal{C}$ and $Jx \in \mathcal{C}$. If $a \ge 0$, this means that $x \in \Theta_a$, which contradicts the hypothesis on x. So we may suppose $a < 0$. We have four cases:
   (a) $|x_1| \le |x_2|$ and $|x_3| \le |x_2|$
   (b) $|x_1| \le |x_2|$ and $|x_3| \ge 2|x_2|$
   (c) $|x_1| \ge 2|x_2|$ and $|x_3| \le |x_2|$
   (d) $|x_1| \ge 2|x_2|$ and $|x_3| \ge 2|x_2|$

Case (a) is impossible since otherwise x would be in Θ_a. If $|x_1| \geq |2x_2|$ then by Equation (1.2), we have

$$0 > a = x_1^2 - 2x_2^2 + x_3^2 \geq 2x_2^2 + x_3^2 \geq 0$$

which is impossible. The cases where $|x_3| \geq 2|x_2|$ are done similarly.

3. Let $\varepsilon \in \{-1, 1\}$ and suppose $x \in \Gamma_a^\varepsilon$. We have

$$\varphi_a(x) = B^{-\varepsilon}x = \begin{pmatrix} 3 & -4\varepsilon & 0 \\ -2\varepsilon & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3x_1 - 4\varepsilon x_2 \\ -2\varepsilon x_1 + 3x_2 \\ x_3 \end{pmatrix}$$

hence

$$\varphi_a(x)_2 = -2\varepsilon x_1 + 3x_2 = x_2 + 2(-\varepsilon x_1 + x_2). \tag{4.1}$$

Note that by definition of Γ_a^ε we have $\varepsilon x_1 x_2 > 0$, hence in particular $x_2 \neq 0$ and we need only consider the case where x_2 is positive and the case where x_2 is negative.

Case 1 . x_2 *is positive.* By definition of Γ_a^ε, ε and x_1 have the same sign. Since $|x_1| > |x_2|$ (by definition of Γ_a^ε), we have

$$-\varepsilon x_1 + x_2 = -|x_1| + x_2 < 0.$$

Hence by Equation (4.1), we have $\varphi_a(x)_2 < x_2$. On the other hand, we have $2x_2 > |x_1|$, hence

$$\varphi_a(x)_2 = -2\varepsilon x_1 + 3x_2 = -2|x_1| + 3x_2 > -x_2$$

and we conclude $|\varphi_a(x)_2| < |x_2|$.

Case 2 . x_2 *is negative.* This case is done similarly and is left to the reader.

□

Proof of Corollary 1.5. It is enough to observe that

$$B^{-1}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \quad B\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix} \quad \text{and} \quad J\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

□

5 Miscellaneous results

We give a list of lemmas that will be used various times till the end of the paper.

Lemma 5.1. *Let $x = (x_1, x_2, x_3) \in \mathbb{Z}^3$. For $|a| \leq 2$, if $x \in \Gamma_a \setminus \Theta_a$ then the sequence (x_1^2, x_2^2, x_3^2) is either strictly increasing or strictly decreasing.*

Proof. Suppose that $x \in \Gamma_a \setminus \Theta_a$ and write Equation (1.2) as

$$(x_3^2 - x_2^2) - (x_2^2 - x_1^2) = a.$$

Cases $a = 1$ and $a = 2$. We have $x_2 \neq 0$ (otherwise $x_1^2 + x_3^2 = a$ and $x \in \Theta_a$). If $x_1^2 = x_2^2$ then $x_3^2 - x_2^2 = a$, which is not possible (if $a = 1$, it would imply $x_2 = 0$). Hence $x_1^2 \neq x_2^2$ and by symmetry we have $x_3^2 \neq x_2^2$.

If $x_1^2 < x_2^2$ then $x_3^2 - x_2^2 = a + (x_2^2 - x_1^2) > a > 0$, hence $x_3^2 > x_2^2$ and the sequence is strictly increasing. If $x_1^2 > x_2^2$ then $x_3^2 - x_2^2 = a + (x_2^2 - x_1^2) < a$, hence $x_3^2 - x_2^2 \leq 1$. Since $x_2 \neq 0$ and $x_2^2 \neq x_3^2$, we deduce that $x_3^2 - x_2^2 < 0$, which implies that the sequence is strictly decreasing.

Cases $a = -1$ **and** $a = -2$**.** We have $x_2^2 \neq 1$ (otherwise $x_3^2 + x_1^2 = a + 2$ and $x \in \Theta_a$) and $x_2 \neq 0$ (otherwise we would have $x_3^2 + x_1^2 = a < 0$). Therefore, we have $x_2^2 \geq 4$ and the difference between x_2^2 and any other square is at least 3 or non-positive. In particular, if $x_2^2 - x_1^2 < -a$ then $x_2^2 < x_1^2$ (note that if $x_1^2 = x_2^2$ then $x_3^2 - x_2^2 = a$, but the difference of two squares cannot be -2, and if $a = -1$ then $x \in \Theta_{-1}$).

If $x_1^2 > x_2^2 + a$ then $x_3^2 - x_2^2 = a + (x_2^2 - x_1^2) < 0$, hence $x_3^2 < x_2^2 < x_1^2$.
If $x_1^2 < x_2^2 + a$ then $x_3^2 - x_2^2 = a + (x_2^2 - x_1^2) > 0$, hence $x_3^2 > x_2^2 > x_1^2$.

Case $a = 0$**.** Note that from Equation (1.2) if $x_i^2 = x_j^2$ for some $i \neq j$, then $x_1^2 = x_2^2 = x_3^2$, in which case $x \in \Theta_0$. If $x_1^2 < x_2^2$ then $0 = x_1^2 - 2x_2^2 + x_3^2 < x_3^2 - x_2^2$ and the sequence is strictly increasing. If $x_1^2 > x_2^2$ then $0 = x_1^2 - 2x_2^2 + x_3^2 > x_3^2 - x_2^2$ and the sequence is strictly decreasing. □

Lemma 5.2. *Let* $\varepsilon = \pm 1$ *and* $x = (x_1, x_2, x_3) \in \mathbb{Z}^3$*. Writing* $(y_1, y_2, x_3) = B^\varepsilon x$*, we have*

1. *if* $\varepsilon x_1 x_2 \geq 0$ *then* $\varepsilon y_1 y_2 \geq 0$*;*
2. *if* $\varepsilon x_1 x_2 > 0$ *then* $\varepsilon y_1 y_2 > 0$*; and*
3. *if* $|x_2| > |x_1|$ *and* $\varepsilon x_1 x_2 < 0$ *then* $\varepsilon y_1 y_2 > 0$*.*

Proof. By definition of B, we have

$$\begin{aligned} \varepsilon y_1 y_2 &= \varepsilon(3x_1 + 4\varepsilon x_2)(2\varepsilon x_1 + 3x_2) \\ &= 6x_1^2 + 17\varepsilon x_1 x_2 + 12x_2^2 \end{aligned}$$

and we can deduce Items 1 and 2. For Item 3, note that

$$\begin{aligned} \varepsilon y_1 y_2 &= 6x_1^2 + 17\varepsilon x_1 x_2 + 12x_2^2 \\ &= 6(x_1 + \varepsilon x_2)^2 + 5\varepsilon x_1 x_2 + 6x_2^2 \\ &= 6(x_1 + \varepsilon x_2)^2 + 5x_2(\varepsilon x_1 + x_2) + x_2^2 \end{aligned}$$

and since $\varepsilon x_1 + x_2$ has the same sign as x_2, we have $5x_2(\varepsilon x_1 + x_2) > 0$. □

Lemma 5.3. *Let* $\varepsilon = \pm 1$ *and* $|a| \leq 2$*. Any strictly increasing sequence* (x_1, x_2, x_3) *(in absolute value) in* Γ_a*, when multiplied by* B^ε*, produces a strictly decreasing sequence (in absolute value)* (y_1, y_2, x_3) *in* Γ_a *satisfying* $\varepsilon y_1 y_2 > 0$*.*

Proof. We first prove that (y_1, y_2, x_3) is not in Θ_a. Since (x_1, x_2, x_3) is strictly increasing in absolute value, we have $|x_2| \geq 1$ and $|x_3| \geq 2$, hence the only cases to check are when $a = 0$, and when $a = 2$ and $(x_1, x_2, x_3) = (0, \pm 1, \pm 2)$. In the latter case, we have $B^\varepsilon(0, \pm 1, \pm 2) = (\pm 4, \pm 3, \pm 2)$ which is not in Θ_2. Suppose for a contradiction that (y_1, y_2, x_3) is in Θ_0 (hence in particular $y_1 = y_2$). Since by definition of B we have $y_1 = 3x_1 + 4\varepsilon x_2$ and $y_2 = 2\varepsilon x_1 + 3x_2$, we obtain $(3 - 2\varepsilon)x_1 = (3 - 4\varepsilon)x_2$, hence

$$1 < \frac{|x_2|}{|x_1|} = \frac{|3 - 2\varepsilon|}{|3 - 4\varepsilon|} \leq 1$$

which is absurd.

Therefore, by Lemma 5.1, it is enough to show that $|y_1| > |y_2|$ and $\varepsilon y_1 y_2 > 0$.

Suppose that $\varepsilon x_1 x_2$ is non-negative. We have

$$|y_1| = |\varepsilon y_1| = |3\varepsilon x_1 + 4x_2| > |2\varepsilon x_1 + 3x_2| = |y_2|$$

where the inequality comes from the fact that εx_1 and x_2 have the same sign. Note also that $\varepsilon y_1 y_2$ is non-negative by Lemma 5.2, and since

$$|y_2| = |2\varepsilon x_1 + 3x_2| > |x_2| > 0,$$

we obtain $\varepsilon y_1 y_2 > 0$.

If $\varepsilon x_1 x_2$ is negative, write $u = x_1 + \varepsilon x_2$. We have

$$\begin{aligned}\varepsilon u y_2 &= 2x_1^2 + 5\varepsilon x_1 x_2 + 3x_2^2 \\ &= 2x_1^2 + 4\varepsilon x_1 x_2 + 2x_2^2 + \varepsilon x_1 x_2 + x_2^2 \\ &= 2(x_1 + \varepsilon x_2)^2 + \varepsilon x_1 x_2 + x_2^2\end{aligned}$$

which is positive, since by hypothesis we have $|x_2| > |x_1|$. Since $y_1 = u + \varepsilon y_2$, we deduce

$$|y_1| = |u + \varepsilon y_2| > |y_2|$$

because u is non-zero (by hypothesis) and because u and εy_2 have the same sign. Note also that $\varepsilon y_1 y_2$ is positive by Lemma 5.2. □

Lemma 5.4. *Let $\varepsilon = \pm 1$. Let $x = (x_1, x_2, x_3) \in \mathbb{Z}^3$ be such that $\varepsilon x_1 x_2 \geq 0$ and $x_1 \neq 0$. For each $n \geq 0$, let u_n and v_n be defined by $(u_n, v_n, x_3) = B^{\varepsilon n} x$. For each $n \geq 0$, we have*

1. $|u_{n+1}| > |u_n|$;
2. $|v_{n+1}| > |v_n|$;
3. $u_n \neq 0$;
4. $\varepsilon u_n v_n \geq 0$.

In particular, $v_n \neq 0$ for each $n \geq 1$. Moreover, if $|a| \leq 2$ and $x \in \Gamma_a$ is strictly decreasing in absolute value (hence $v_n \neq 0$), then (u_n, v_n, x_3) is strictly decreasing in absolute value (it is false in general if x does not satisfy the hypothesis $\varepsilon x_1 x_2 \geq 0$).

Proof. Note that the lemma is trivial for $n = 0$. Suppose that the lemma holds for some integer $n \geq 0$. Since $\varepsilon u_n v_n \geq 0$ and $u_n \neq 0$ we have

$$|u_{n+1}| = |3u_n + 4\varepsilon v_n| > |u_n|$$

and

$$|v_{n+1}| = |2\varepsilon u_n + 3v_n| > |v_n|$$

(where the equalities come from the definition of B). Hence also $u_{n+1} \neq 0$ and

$$\begin{aligned}\varepsilon u_{n+1} v_{n+1} &= \varepsilon(3u_n + 4\varepsilon v_n)(2\varepsilon u_n + 3v_n) \\ &= 6u_n^2 + 12v_n^2 + 17\varepsilon u_n v_n\end{aligned}$$

is non-negative.

We now prove the last statement of the lemma. If $n = 0$ there is nothing to prove, so we assume $n \geq 1$. By Lemma 5.1, it is enough to prove that (u_n, v_n, x_3) is not in Θ_a (the point is that x_3 does not change as n varies and $|x_3|$ remains the minimum of the sequence of absolute values).

Since $n \geq 1$, we have both $u_n \neq 0$ and $v_n \neq 0$. Hence the only possibilities for (u_n, v_n, x_3) to be in Θ_a are when $a = -1$ and $(u_n, v_n, x_3) = (\pm 1, \pm 1, 0)$, or $a = 0$, or $a = 2$ and $(u_n, v_n, x_3) = (\pm 2, \pm 1, 0)$. By Item 2, if $v_n = \pm 1$ then $n = 1$. We have $B^{\varepsilon}(x_1, x_2, x_3) = (3x_1 + 4\varepsilon x_2, 2\varepsilon x_1 + 3x_2, x_3)$. When $a = 2$, this leads to $3x_1 + 4\varepsilon x_2 = \pm 2$, which is impossible since $3x_1$ and $4\varepsilon x_2$ are of the same sign by hypothesis. An analogous argument discards the case $a = -1$. If $a = 0$ then by Item 1 we have $|u_n| \geq |u_1| > |u_0| = |x_1| > |x_3|$ since the initial sequence is supposed to be strictly decreasing. □

Lemma 5.5. *If*

$$w = B^{n_k} J \ldots B^{n_1} J$$

is an element of H, where $k \geq 1$ and each n_i is a non-zero integer, then its third column is strictly decreasing in absolute value and the entry w_{23} of the matrix w at line 2 and column 3 is distinct from 0.

Proof. We prove by induction on the right subwords

$$W^s = B^{n_s} J \dots B^{n_1} J^r$$

of w that the third column of each W^s is strictly decreasing in absolute value and that the entry W^s_{23} of the matrix W^s at line 2 and column 3 is distinct from 0.

Suppose that $s = 1$. Let ε be 1 if n_1 is positive and -1 otherwise. By Lemma 5.4, taking for x the third column of the matrix J, we need only prove that the third column of W^1 is strictly decreasing in absolute value. Let u_n and v_n be like in Lemma 5.4. Since $n_1 \geq 1$, we have $v_{n_1} \neq 0$ (by Lemma 5.4), hence the third column $(u_{n_1}, v_{n_1}, 0)$ of W^1 is not in Θ_1 and we can apply Lemma 5.1, which implies that $(u_{n_1}, v_{n_1}, 0)$ is strictly decreasing in absolute value.

Suppose that the property holds up to $s-1$. Hence by hypothesis of induction, the third column of W^{s-1} is an element (x_3, x_2, x_1) of Γ_1, such that $x_2 \neq 0$ and which is strictly decreasing in absolute value. When multiplied by J, it becomes a strictly increasing sequence (in absolute value) (x_1, x_2, x_3). Therefore, by Lemma 5.3, when the latter is multiplied by B^ε, it gives a strictly decreasing (in absolute value) sequence (y_1, y_2, x_3) in Γ_1 such that $\varepsilon y_1 y_2$ is positive. By Lemma 5.4, taking for x the third column (y_1, y_2, x_3) of the matrix $B^\varepsilon J W^{s-1}$, we need only prove that the third column of W^s is strictly decreasing in absolute value. If $n_s = \pm 1$, then we have nothing more to prove. If $n_s \geq 2$, letting u_n and v_n be like in Lemma 5.4, we have $v_{n_s - 1} \neq 0$ and we can conclude that $(u_{n_s-1}, v_{n_s-1}, x_3)$ is strictly decreasing in absolute value. □

We finish this section by a folklore Lemma.

Lemma 5.6. *If $y = (y_1, y_2, \dots, y_N)$ is a non-trivial Büchi sequence of length N which is increasing in absolute value then, for each index $n \geq 2$, we have*

$$|y_{n+1}| - |y_n| < |y_n| - |y_{n-1}|. \tag{5.1}$$

Proof. If for some index $n \geq 2$ we have $|y_{n+1}| - |y_n| \geq |y_n| - |y_{n-1}|$ then $|y_{n+1}| \geq 2|y_n| - |y_{n-1}|$, hence

$$2 - y_{n-1}^2 + 2y_n^2 = y_{n+1}^2 \geq 4y_{n^2} - 4|y_n y_{n-1}| + y_{n-1}^2$$

and we get

$$2 \geq 2y_n^2 - 4|y_n y_{n-1}| + 2y_{n-1}^2$$

hence

$$1 \geq (|y_n| - |y_{n-1}|)^2$$

which implies that the sequence is trivial. □

6 Presentation of the group H

Theorem 1.6 is an easy corollary of Lemma 5.5. We consider an arbitrary element of H

$$w = J^\ell B^{n_k} J \dots B^{n_2} J B^{n_1} J^r$$

where $k \geq 1$, each n_i is a non-zero integer, and ℓ and r are 0 or 1. We will prove that w is *not* the identity matrix and the theorem will follow (since the only non-empty word that we are missing is J which is distinct from I).

Note that if $\ell = 1$ then it is enough to show that JwJ is not the identity, and if $\ell = r = 0$ then it is enough to show that $B^{n_k} w B^{-n_k}$ is not the identity. So, without loss of generality, we can assume $\ell = 0$ and $r = 1$, and conclude with Lemma 5.5.

7 Proof of Theorem 1.7

By Corollary 1.5, we need only prove the *unicity* part of the theorem.

Definition 7.1. If $M = J^\ell B^{n_k} J \dots B^{n_2} J B^{n_1} J^r$, where $k \geq 1$, each n_i is a non-zero integer, and ℓ and r are 0 or 1, then we will call k the *length of* M. Elements of H of *length* 0 are I and J. We will refer to $(\ell, n_k, \dots, n_1, r)$ as to the *sequence of powers associated to* M.

Next Lemma is a corollary of Lemma 5.5 which we already used in order to find the presentation of H.

Lemma 7.2. *If $M \in H$ is such that $M_{23} = 0$ then either $M = I$ or $M = J$ or $M = B^n$ or $M = JB^n$ for some $n \in \mathbb{Z}$.*

Proof. Suppose that M is neither I, nor J, and neither of the form B^n nor JB^n. Hence in particular M has length at least 1 and can be written as

$$M = J^\ell B^{n_k} J \dots B^{n_2} J B^{n_1} J^r$$

for some $k \geq 1$ and where each n_i is a non-zero integer, and ℓ and r are 0 or 1. We want to prove that M_{23} is non-zero.

If $\ell = 0$ and $r = 1$ then we conclude by Lemma 5.5. Also if $\ell = 1$ and $r = 1$ then $(JM)_{23}$ is non-zero by Lemma 5.5, hence M_{23} is non-zero. So we may suppose that $r = 0$.

Since the only words of length 1 with $r = 0$ are of the form B^n or JB^n, we may suppose that the length of M is at least 2. Let $M_0 \in H$ be such that $M = M_0 B^{n_2} J B^{n_1}$. By Lemma 5.5, we have $(M_0 B^{n_2} J)_{23} \neq 0$. We can conclude that M_{23} is non-zero because multiplying by B on the right does not affect the third column. □

Next lemma resumes some basic properties of the matrix B.

Lemma 7.3. *The characteristic polynomial of B is $x^3 - 7x^2 + 7x - 1$, its eigen values are $2\sqrt{2} + 3$, $-2\sqrt{2} + 3$ and 1, and*

$$\begin{pmatrix} \sqrt{2} & -\sqrt{2} & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is a matrix of eigen vectors. Hence for any $n \in \mathbb{Z}$ we have

$$B^n = \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{2}(\bar{\alpha}^n + \alpha^n) & 2(-\bar{\alpha}^n + \alpha^n) & 0 \\ -\bar{\alpha}^n + \alpha^n & \sqrt{2}(\bar{\alpha}^n + \alpha^n) & 0 \\ 0 & 0 & 2\sqrt{2} \end{pmatrix}$$

where $\alpha = 2\sqrt{2} + 3$ and $\bar{\alpha} = \alpha^{-1}$ is the conjugate of α in $\mathbb{Z}[\sqrt{2}]$. Moreover, each entry (i, j) in B^n, with $i, j \in \{1, 2\}$, satisfies the recurrence relation $B^n_{i,j} = 6B^{n-1}_{i,j} - B^{n-2}_{i,j}$ (the initial values are given by the identity matrix and B at the corresponding entry).

We believe that the recurrence relation described above could be very useful to solve Problems A and B (see Section 9). For the purposes of this section, we will only need the following:

Corollary 7.4. *The matrices B^n and JB^n, for $n \in \mathbb{Z} \setminus \{0\}$, have second row distinct from $(0, \pm 1, 0)$, and the diagonal entries are positive integers.*

Proof. Observe that both α and $\bar{\alpha}$ are positive real numbers. □

Definition 7.5. A sequence in Γ_2 is *odd* if it is in the orbit of one of $(\pm 1, 0, \pm 1)$ and it is *even* if it is in the orbit of one of $(\pm 2, 1, 0)$.

Lemma 7.6. *If a sequence* $(x_1, x_2, x_3) \in \Gamma_2$ *is odd then* x_1 *and* x_3 *are odd, and* x_2 *is even. If it is even, then* x_1 *and* x_3 *are even and* x_1 *is odd.*

Proof. If x_1 and x_3 are odd and x_2 is even, then $3x_1 + 4x_2$ is odd and $2x_1 + 3x_2$ is even, hence any odd sequence in Γ_2 satisfies the desired property. The case of even sequences is done similarly. □

Next lemma finishes the proof of the theorem.

Lemma 7.7. *Let* $M \in H$ *and* $\delta, \delta' \in \Delta_2$ *such that* $M\delta = \delta'$*. If* δ *is odd then* M *is either* I *or* J *(in the latter case,* δ *must be* $(1, 0, 1)$ *or* $(-1, 0, -1)$*). If* δ *is even, then* M *is the identity. In all cases we have* $\delta = \delta'$*.*

Proof. Write $M = (m_{ij})$ and suppose first that δ is odd, i.e. $\delta = (\varepsilon_1, 0, \varepsilon_2)$ for some $\varepsilon_1, \varepsilon_2 \in \{\pm 1\}$. Since $M\delta = \delta'$, we have $\varepsilon_1 m_{21} + \varepsilon_2 m_{23} = 0$. Since the second row of M is in Ω_1 (see Theorem 1.2), we have

$$-2m_{21}^2 + m_{22}^2 - 2m_{23}^2 = 1$$

hence

$$(m_{22} - 2m_{21})(m_{22} + 2m_{21}) = m_{22}^2 - 4m_{21}^2 = 1.$$

We have then $m_{22} - 2m_{21} = m_{22} + 2m_{21}$, hence $m_{21} = 0$ and $m_{22}^2 = 1$. So the second row of M is $(0, \pm 1, 0)$ and we conclude by Lemma 7.2 and Corollary 7.4.

Suppose now that δ is even, i.e. $\delta = (2\varepsilon, 1, 0)$ for some $\varepsilon \in \{\pm 1\}$. Since δ' is in the orbit of δ, it is also even by Lemma 7.6, so $\delta' = (2\varepsilon', 1, 0)$ for some $\varepsilon' \in \{\pm 1\}$. Since $M\delta = \delta'$, we have $2\varepsilon m_{31} + m_{32} = 0$. Since the third row is in Ω_{-2} (see Theorem 1.2), we have

$$-2m_{31}^2 + m_{32}^2 - 2m_{33}^2 = -2$$

hence $2m_{31}^2 - 2m_{33}^2 = -2$, which implies $m_{31} = m_{32} = 0$ and $m_{33} = \pm 1$. Since the third column is in Γ_1, we have $m_{13}^2 - 2m_{23}^2 + m_{33}^2 = 1$, hence $m_{13}^2 - 2m_{23}^2 = 0$, which implies $m_{13} = m_{23} = 0$. By Lemma 7.2, the only possibilities for M are I, B^n or JB^n for some $n \in \mathbb{Z}$. Hence in particular, we can assume that all m_{ii} are positive by Corollary 7.4 (hence $m_{33} = 1$).

On the other hand, we also have $2\varepsilon m_{11} + m_{12} = 2\varepsilon'$. Since the second row is in Ω_1, we have $-2m_{21}^2 + m_{22}^2 - 2m_{23}^2 = 1$, hence $-2m_{21}^2 + m_{22}^2 = 1$, hence

$$-(1 - m_{22})^2 + 2m_{22}^2 = 2$$

and we finally obtain two solutions for m_{22}, which are 1, in which case $M = I$; or -3, which is impossible. □

8 Congruences modulo 8

Next theorem shows that in order to know in which orbit a length 3 Büchi sequence is, it is enough to consider the sequence modulo 8 ('congruent' means 'congruent modulo 8' in this section).

Theorem 8.1. *A Büchi sequence* $x = (x_1, x_2, x_3)$ *is in the orbit of:*

1. $(1, 0, 1)$ *if and only if both* x_1 *and* x_3 *are congruent to* 1 *or* 3*;*

2. $(-1, 0, -1)$ *if and only if both* x_1 *and* x_3 *are congruent to* -1 *or* -3*;*

3. $(-1, 0, 1)$ *if and only if* x_1 *is congruent to* -1 *or* -3 *and* x_3 *is congruent to* 1 *or* 3*, or* x_3 *is congruent to* -1 *or* -3 *and* x_1 *is congruent to* 1 *or* 3*;*

4. $(2, 1, 0)$ *if and only if either* x_1 *or* x_3 *is congruent to* 2*; and*

5. $(-2,1,0)$ *if and only if either* x_1 *or* x_3 *is congruent to* -2.

Proof. Recall that

$$Bx = \begin{pmatrix} 3x_1 + 4x_2 \\ 2x_1 + 3x_2 \\ x_3 \end{pmatrix} \qquad \text{and} \qquad B^{-1}x = \begin{pmatrix} 3x_1 - 4x_2 \\ -2x_1 + 3x_2 \\ x_3 \end{pmatrix}.$$

Suppose first that x is an odd sequence. Since x_2 is even (see Lemma 7.6), $3x_1 \pm 4x_2$ is congruent to $3x_1$. Hence, if x_1 is congruent to 1 or 3 then $3x_1 \pm 4x_2$ is also congruent to 1 or 3. Similarly, if x_1 is congruent to -1 or -3 then $3x_1 \pm 4x_2$ is also congruent to -1 or -3. From this observations and the fact that multiplying by J interchanges x_1 and x_3, it is easy to conclude for Items 1, 2 and 3 of the Theorem.

If x is an even sequence then x_2 is odd and $3x_1 + 4x_2$ is congruent to $3x_1 + 4$. So if x_1 is congruent to 2 then also $3x_1 + 4x_2$ is congruent to 2, and if x_1 is congruent to -2 then also $3x_1 + 4x_2$ is congruent to -2. This allows us to conclude for Items 4 and 5. □

Next Lemma says that Büchi's problem has a positive answer for $\mathbb{Z}/8\mathbb{Z}$ (Hensley [H2] solved Büchi's problem modulo any power of a prime, but did not try to find optimal lower bounds for the length of non-trivial sequences).

Lemma 8.2. *Modulo* 8*, all Büchi sequences of length* 3 *are trivial.*

Proof. Let $x = (x_1, x_2, x_3)$ be a Büchi sequence modulo 8. Squares are 0, 1 and 4. If $x_1^2 = 0$ then $-2x_2^2 + x_3^2 = 2$, hence $x_2^2 = 1$ and $x_3^2 = 4$. Therefore, the sequence (x_1^2, x_2^2, x_3^2) is a sequence of consecutive squares, which implies that x is trivial. If $x_1^2 = 1$ then $-2x_2^2 + x_3^2 = 1$, hence $x_2^2 = 0$ or $x_2^2 = 4$. If $x_2^2 = 0$ then $x_3^2 = 1$ and we obtain a sequence of consecutive squares. If $x_2^2 = 4$ then $x_3^2 = 1$, but again the sequence $(1,4,1) = (1^2, 2^2, 3^2)$ is a sequence of consecutive squares. □

Remark 8.3. If $x = (x_1, x_2, x_3)$ is an even Büchi sequence and if for example x_1 is congruent to ± 2, then by Lemma 8.2 its sequence of squares is either of the form $(2^2, 3^2, 4^2)$ or $(2^2, 1^2, 0^2)$, hence x_3 is congruent to 0 or 4. Unfortunately, this argument does not give any information for odd sequences.

Next corollaries are the key points of our strategy to solve Büchi's problem (see Section 9).

Corollary 8.4. *Given a length* 5 *Büchi sequence* $(x_1, \dots, x_5)$*, after changing the signs of* x_1*,* x_3 *or* x_5 *if necessary,* (x_1, x_2, x_3) *and* (x_3, x_4, x_5) *are both in the orbit of*

1. $(-1,0,1)$ *if* x_1 *is odd; and*
2. $(2,1,0)$ *if* x_1 *is even.*

Proof. It is immediate from Theorem 8.1. □

Before stating next corollary, let us first introduce two definitions.

Definition 8.5. A Büchi sequence $(x_1, \dots, x_M)$ is *odd* if (x_1, x_2, x_3) is odd and it is *even* if not.

Definition 8.6. We will call a length 5 sequence of integers $x = (x_1, \dots, x_5)$ *canonical* if it satisfies

1. x_1 and x_5 are congruent to 2; and
2. either x_4 is congruent to 1 or -3, and x_2 is congruent to -1 or 3, or x_4 is congruent to -1 or 3, and x_2 is congruent to 1 or -3.

Note that in the definition above we do not require the sequence to be a Büchi sequence.

Corollary 8.7. *Given a length* 8 *Büchi sequence* $y = (y_1, \dots, y_8)$*, after changing the signs of the* y_i *if necessary, there exists* $1 \le j \le 4$ *such that* $(y_j, \dots, y_{j+4})$ *is canonical.*

Proof. Let $z = (z_1, \dots, z_7)$ be the (unique) even length 7 subsequence of y. Let $k \in \{1, 3\}$ be such that z_k is congruent to ± 2 (such a k exists by Theorem 8.1). Write $x = (x_1, \dots, x_5) = (z_k, \dots, z_{k+4})$ (so the index j of the statement can be chosen to be k if $z_1 = y_1$ and $k+1$ if $z_1 = y_2$).

Since $x_1 = z_k$ is congruent to ± 2, by Remark 8.3, x_3 is congruent to 0 or 4, and by Theorem 8.1, x_5 is congruent to ± 2. Also by Theorem 8.1, both x_2 and x_4 are congruent to ± 1 or ± 3. So we can obtain the desired sequence by multiplying x_1, x_2 and x_5 by -1 if necessary. □

Corollary 8.8. *If all canonical Büchi sequences are trivial then all length* 8 *Büchi sequences are trivial.*

Proof. Let y be a length 8 Büchi sequence and x be a canonical subsequence of y (it exists by Corollary 8.7). Since x is trivial by hypothesis, also y is trivial (indeed it is easy to see that if there are two consecutive terms x_i and x_{i+1} in a Büchi sequence such that $|x_i| = |x_{i+1}| \pm 1$ then the sequence is trivial). □

9 A strategy for Büchi's Problem

Let $x = (x_1, \dots, x_5)$ be a length 5 Büchi sequence. By changing the signs of x_1, x_3 or x_5 if necessary, we may suppose that (x_1, x_2, x_3) and (x_3, x_4, x_5) are both in the orbit of $(2, 1, 0)$, or both in the orbit of $(-1, 0, 1)$ (see Corollary 8.4). By Theorem 1.7, we know that there exist unique matrices M_1, M_2 and M_3 and unique $\delta, \delta' \in \Delta_2$ such that

$$\begin{cases} (x_1, x_2, x_3) = M_1 \delta \\ (x_2, x_3, x_4) = M_2 \delta' \\ (x_3, x_4, x_5) = M_3 \delta \end{cases} \tag{9.1}$$

and if we write $M_x = J M_3 M_1^{-1}$ then we have

$$M_x \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_5 \\ x_4 \\ x_3 \end{pmatrix}. \tag{9.2}$$

Note that the matrix M_x is uniquely determined by x *once the signs of the* x_i *have been chosen.*

Lemma 9.1. *If* $M_x = B$ *or* B^{-1} *then the sequence* x *is trivial.*

Proof. If $M_x = B$ or B^{-1} then we have $2x_1 \pm 3x_2 = x_4$, hence

$$2 = x_4^2 - 2x_3^2 + x_2^2 = (2x_1 \pm 3x_2)^2 - 2x_3^2 + x_2^2 = 4x_1^2 \pm 12x_1x_2 + 10x_2^2 - 2x_3^2$$

and since $x_3^2 - 2x_2^2 + x_1^2 = 2$, this gives

$$2 = 4x_1^2 \pm 12x_1x_2 + 10x_2^2 - 2(2 - x_1^2 + 2x_2^2) = 6x_1^2 \pm 12x_1x_2 + 6x_2^2 - 4$$

hence $x_1^2 \pm 2x_1x_2 + x_2^2 = 1$, which implies that $x_1 \pm x_2 = \varepsilon$, for some $\varepsilon \in \{-1, 1\}$. Writing $\nu = -\varepsilon x_1$, one conclude easily that for each i we have $x_i^2 = (\nu + i - 1)^2$, so the sequence x is a trivial Büchi sequence. □

We may write $\xi_1 = (x_1, x_2, x_3)$ and $\xi_2 = (x_5, x_4, x_3)$, so that we have

$$M_x \xi_1 = \xi_2.$$

Also for any sequence y, we will denote by $|y|$ the sequence of its absolute values.

In order to prove that there is no non-trivial Büchi sequence of length 5, one strategy is to try to solve the following problem by induction on n.

Problem A. Is it true that for all $n \geq 0$, sequences x whose matrix M_x has length n are trivial?

Next Lemma shows that Problem A has a positive answer for $n \leq 1$.

Theorem 9.2. *If M_x has length ≤ 1 then x is trivial.*

Proof. We will assume that x is non-trivial and obtain a contradiction when M_x has length 0 or 1.

By Lemma 5.1, since x is non-trivial, the sequence $|x|$ is either strictly increasing or strictly decreasing. Suppose first that it is strictly increasing.

If the length of M were 0 then we would have $M = I$ or J, hence $x_1 = x_5$ or, respectively, $x_1 = x_3$, which would give a contradiction in both cases.

For the sake of contradiction, assume that the length of M is 1, so that M has one of the following forms: $B^{\varepsilon n}$, $JB^{\varepsilon n}$, $JB^{\varepsilon n}J$ or $B^{\varepsilon n}J$, where $n \geq 1$ and $\varepsilon = \pm 1$.

Case $M = B^{\varepsilon n}J$. We have $(x_5, x_4, x_3) = B^{\varepsilon n}J\xi_1 = B^{\varepsilon n}(x_3, x_2, x_1)$, hence $x_1 = x_3$, which is impossible.

Case $M = JB^{\varepsilon n}J$. We have $(x_3, x_4, x_5) = J\xi_2 = B^{\varepsilon n}J\xi_1 = B^{\varepsilon n}(x_3, x_2, x_1)$, hence $x_1 = x_5$, which is impossible.

Case $M = JB^{\varepsilon n}$. Since $|\xi_1|$ is strictly increasing, the sequence (y_1, y_2, x_3) defined by $B^{\varepsilon}\xi_1$ is strictly decreasing in absolute value and satisfies $\varepsilon y_1 y_2 > 0$ (see Lemma 5.3). By Lemma 5.4, $B^{\varepsilon(n-1)}B^{\varepsilon}\xi_1 = J\xi_2 = (x_3, x_4, x_5)$ is strictly decreasing in absolute value, which is impossible.

Case $M = B^{\varepsilon n}$. Since x is assumed to be non-trivial, we have $n > 1$ by Lemma 9.1. We first prove that if (y_1, y_2, x_3) is defined by $B^{\varepsilon}\xi_1$ then $|y_1| > |x_5|$. We have

$$\begin{aligned} |y_1| = |3x_1 + 4\varepsilon x_2| \geq 4|x_2| - 3|x_1| &= 3(|x_2| - |x_1|) + |x_2| \\ &> 3(|x_3| - |x_2|) + |x_2| = 2(|x_3| - |x_2|) + |x_3| \\ &> 2(|x_4| - |x_3|) + |x_3| = (|x_4| - |x_3|) + |x_4| \\ &> (|x_5| - |x_4|) + |x_4| = |x_5| \end{aligned}$$

where the strict inequalities come from Lemma 5.6. By Lemma 5.3, the sequence (y_1, y_2, x_3) is strictly decreasing in absolute value and satisfies $\varepsilon y_1 y_2 > 0$, hence applying Lemma 5.4 repeatedly $(n-1)$ times, the sequence

$$(x_5, x_4, x_3) = M\xi_1 = B^{\varepsilon(n-1)}B\xi_1$$

satisfies $|x_5| > |x_5|$, which is absurd. So the lemma is proven for x strictly increasing in absolute value.

Suppose now that $|x|$ is strictly decreasing and consider $\bar{x} = (x_5, \ldots, x_1)$. There exists a unique matrix $M_{\bar{x}}$ such that $M_{\bar{x}}(x_5, x_4, x_3) = (x_1, x_2, x_3)$, hence $M_{\bar{x}}^{-1}(x_1, x_2, x_3) = (x_5, x_4, x_3)$. Therefore, we have $M_{\bar{x}}^{-1} = M_x$ and since $|\bar{x}|$ is strictly increasing, we know from the study above that $M_{\bar{x}}$, hence also $M_{\bar{x}}^{-1} = M_x$, cannot have length ≤ 1 if x is non-trivial. □

Remark 9.3. Suppose that we want to prove that there is no non-trivial Büchi sequences of length 6. Since in a Büchi sequence of length 6, there is exactly one odd subsequence of length 5 and one even subsequence of length 5 (see Definition 8.5), it is enough to show that there is no odd sequence of length 5 or that there is no even sequence of length 5. Therefore, it would be enough to solve Problem A for $n \geq 2$ and assuming, for example, that x is in the orbit of $(2,1,0)$.

We will finish this section by presenting a strategy to try to prove that all Büchi sequences of length 8 are trivial.

The reciprocal of Lemma 9.1 is not true in general. Indeed, there are counter-examples for both odd and even sequences. For example with $x = (-1,2,3,-4,5)$, we have $\delta = (-1,0,1)$, $M_1 = JBJ$ and $M_3 = JB^{-1}JB^{-1}J$, hence $M_x = B^{-1}JB^{-2}J$. With $x = (2,3,4,-5,-6)$, we have $\delta = (2,1,0)$, $M_1 = JBJ$ and $M_3 = JBJB^{-1}JB^{-1}$, hence $M_x = BJB^{-1}JB^{-1}JB^{-1}J$.

Lemma 9.4. *Assume that x is canonical (as defined in 8.6). If x is trivial then $M_x = B$ or B^{-1}.*

Proof. Since x is trivial, there exists an integer $n \in \mathbb{Z}$ such that $x_i = \varepsilon_i(n+i)$, where $\varepsilon_i \in \{-1,1\}$ for each $i = 1,\ldots,5$. Writing $x_1 = 8m+2$, we have

$$n = \varepsilon_1(8m+2) - 1$$

hence

$$x_5 = \varepsilon_5(\varepsilon_1(8m+2) + 4)$$

and since x_5 is by hypothesis congruent to 2 modulo 8, we have $\varepsilon_5\varepsilon_1 = -1$. Also we have

$$x_4 = \varepsilon_4(\varepsilon_1(8m+2) + 3) \qquad \text{and} \qquad x_2 = \varepsilon_2(\varepsilon_1(8m+2) + 1)$$

hence

- x_4 is congruent to 1 or -3 if and only if $\varepsilon_4 = 1$; and
- x_2 is congruent to -1 or 3 if and only if $\varepsilon_2 = 1$.

Since the sequence is canonical, we have $\varepsilon_2 = \varepsilon_4$.

Writing $\varepsilon = -\varepsilon_1\varepsilon_2$, we have

$$B^\varepsilon \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = B^\varepsilon \begin{pmatrix} \varepsilon_1(n+1) \\ \varepsilon_2(n+2) \\ x_3 \end{pmatrix} = \begin{pmatrix} 3\varepsilon_1(n+1) + 4\varepsilon\varepsilon_2(n+2) \\ 2\varepsilon\varepsilon_1(n+1) + 3\varepsilon_2(n+2) \\ x_3 \end{pmatrix}$$

hence

$$B^\varepsilon \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_1(3(n+1) - 4(n+2)) \\ \varepsilon_2(-2(n+1) + 3(n+2)) \\ x_3 \end{pmatrix} = \begin{pmatrix} \varepsilon_1(-n-5) \\ \varepsilon_2(n+4) \\ x_3 \end{pmatrix}$$

and we can conclude since $\varepsilon_1 = -\varepsilon_5$ and $\varepsilon_2 = \varepsilon_4$. □

Problem B. Let $x = (x_1,\ldots,x_5)$ be a canonical sequence. Suppose that there exist matrices M_1, M_2 and M_3 in H such that $(x_1,x_2,x_3) = M_1(2,1,0)$, $(x_2,x_3,x_4) = M_2\delta$ and $(x_3,x_4,x_5) = M_3(2,1,0)$, where δ is $(\pm 1, 0, \pm 1)$. Is it the case that $M_x = JM_3M_1^{-1}$ is necessarily either B or B^{-1}.

Theorem 9.5. *If Problem B has a positive answer then there are no non-trivial Büchi sequences of length* 8*. If there are no non-trivial Büchi sequences of length* 5 *then Problem B has a positive answer.*

Proof. Suppose that Problem B has a positive answer and let y be a Büchi sequence of length 8. By Corollary 8.8 there exists a canonical Büchi subsequence x of y. By Theorem 1.7 there exist matrices M_1, M_2 and M_3 in H satisfying the hypothesis of Problem B. Hence M_x is either B or B^{-1}. By Lemma 9.1, x is a trivial sequence, hence also y is a trivial sequence.

Suppose that there are no non-trivial Büchi sequences of length 5. In particular, there are no non-trivial canonical Büchi sequences. Hence all canonical Büchi sequences are trivial. By Lemma 9.4, this implies that all canonical Büchi sequences x are such that M_x is B or B^{-1}, and Problem B has a positive answer. □

Acknowledgements

We thank J. Browkin, H. Pasten and T. Pheidas for several discussions during the preparation of this work, and the referee for his useful comments.

References

[BB] J. Browkin and J. Brzeziński, *On sequences of squares with constant second differences*, Canad. Math. Bull. **49-4** (2006), 481–491.

[C] J. W. S. Cassels, *Rational Quadratic Forms*, Dover (2008).

[D] M. Davis, *Hilbert's tenth problem is unsolvable*, American Mathematical Monthly **80** (1973), 233–269.

[H1] D. Hensley, *Sequences of squares with second difference of two and a problem of logic*, unpublished, 1980-1983.

[H2] —– *Sequences of squares with second difference of two and a conjecture of Büchi*, unpublished, 1980-1983.

[L] L. Lipshitz, *Quadratic forms, the five square problem, and diophantine equations*, in The collected works of J. Richard Büchi (S. MacLane and Dirk Siefkes, eds.) Springer (1990), 677–680.

[M] Y. Matiyasevic, *Enumerable sets are diophantine*, Dokladii Akademii Nauk SSSR **191**, 279-282 (1970); English translation. Soviet Mathematics Doklady **11** (1970), 354–358.

[Pa2] H. Pasten, *Representation of squares by monic second degree polynomials in the field of p-adic meromorphic functions*, arXiv:1003.1969.

[PPV] H. Pasten, T. Pheidas, X. Vidaux *A survey on Büchi's problem : new presentations and open problems*, to appear in Zapiski POMI, Steklov Institute of Mathematics. Published online http://www.pdmi.ras.ru/znsl/2010/v377.html

[S] W. Sierpiński, *250 problems in elementary number theory*, Elsevier, 1970.

[V] P. Vojta, *Diagonal quadratic forms and Hilbert's Tenth Problem*, Contemporary Mathematics **270** (2000), 261–274 .