Proc. London Math. Soc. (3) 88 (2004) 42 – 62 © 2004 London Mathematical Society
DOI: 10.1112/S002461150301428X

APPROXIMATION TO REAL NUMBERS BY CUBIC ALGEBRAIC INTEGERS I

DAMIEN ROY

1. *Introduction*

The study of approximation to a real number by algebraic numbers of bounded degree started with a paper of E. Wirsing [**10**] in 1961. Motivated by this, H. Davenport and W. M. Schmidt considered in [**5**] the analogous inhomogeneous problem of approximation to a real number by algebraic integers of bounded degree. They proved a result that is optimal for degree 2 and a general result which is valid for any degree. We shall be concerned here with the case of degree 3, the first case for which the optimal exponent of approximation is not known.

To state their results, we define the height $H(P)$ of a polynomial $P \in \mathbb{R}[T]$ to be the maximum of the absolute values of its coefficients, and define the height $H(\alpha)$ of an algebraic number α to be the height of its irreducible polynomial over $\mathbb{Z}$. We also denote by $\gamma = \frac{1}{2}(1+\sqrt{5})$ the golden ratio.

Davenport and Schmidt showed that, for any real number ξ which is neither rational nor quadratic irrational, there exist a positive constant c depending only on ξ and arbitrarily large real numbers X such that the inequalities

$$|x_0| \leqslant X, \quad |x_0\xi - x_1| \leqslant cX^{-1/\gamma}, \quad |x_0\xi^2 - x_2| \leqslant cX^{-1/\gamma} \tag{1.1}$$

have no solutions in integers x_0, x_1, x_2, not all zero (see [**5**, Theorem 1a]). By an argument of duality based on geometry of numbers, they deduced from this that, for such a real number ξ, there exist another constant $c > 0$ and infinitely many algebraic integers α of degree at most 3 which satisfy

$$0 < |\xi - \alpha| \leqslant cH(\alpha)^{-\gamma^2} \tag{1.2}$$

(see [**5**, Theorem 1]).

At first, it would be natural to expect that the first statement holds with any exponent greater than $\frac{1}{2}$ in place of $1/\gamma \simeq 0.618$ in the inequalities (1.1), since this ensures that the volume of the convex body defined by these inequalities tends to zero as X tends to infinity. However, this is not the case, and we will show in fact that the conditions (1.1) in this statement are optimal (see also [**8**] for an announcement).

THEOREM 1.1. *There exists a real number ξ which is neither rational nor quadratic irrational and which has the property that, for suitable constant $c > 0$,*

Received 19 July 2002; revised 5 December 2002.

2000 *Mathematics Subject Classification* 11J04 (primary), 11J13, 11J82 (secondary).

Work partly supported by NSERC and CICMA.

the inequalities (1.1) *have a non-zero solution* $(x_0, x_1, x_2) \in \mathbb{Z}^3$ *for any real number* $X \geqslant 1$. *Any such number is transcendental over* $\mathbb{Q}$ *and the set of these real numbers is countable.*

Let us call *extremal* any real number that has the property stated in Theorem 1.1. For such a number ξ, solutions to (1.1) provide simultaneous approximations of ξ and ξ^2 by rational numbers x_1/x_0 and x_2/x_0 with the same denominator (assuming $x_0 \neq 0$). The fact that, for any sufficiently large X, these are sharper than expected from an application of the Dirichlet box principle (see, for example, [**9**, Chapter II, Theorem 1A]) reminds one of Cassels' counter-example in questions of algebraic independence (see [**3**, Chapter V, Theorem XIV]). Another situation, even closer to that of Cassels, where these numbers satisfy estimates that are better than expected from the box principle for all sufficiently large values of the parameter X is the following (compare with [**9**, Chapter II, Theorem 1B]).

THEOREM 1.2. *Let* ξ *be an extremal real number. There exists a constant* $c > 0$ *depending only on* ξ *such that for any real number* $X \geqslant 1$ *the inequalities*

$$|x_0| \leqslant X, \quad |x_1| \leqslant X, \quad |x_0\xi^2 + x_1\xi + x_2| \leqslant cX^{-\gamma^2} \tag{1.3}$$

have a non-zero solution $(x_0, x_1, x_2) \in \mathbb{Z}^3$.

Indeed, since $\gamma^2 \simeq 2.618 > 2$, the volume of the convex body of $\mathbb{R}^3$ defined by (1.3) tends to zero as X tends to infinity, a situation where we do not expect in principle to find a non-zero solution for each sufficiently large X.

Examples of extremal real numbers are the numbers $\xi_{a,b}$ whose continued fraction expansion $\xi_{a,b} = [0, a, b, a, a, b, a, b, \ldots]$ is given by the Fibonacci word on two distinct positive integers a and b (the infinite word $abaabab\ldots$ starting with a which is the fixed point of the substitution sending a to ab and b to a). We do not know if, like these numbers $\xi_{a,b}$, all extremal real numbers have bounded partial quotients but they satisfy the following weaker measure of approximation by rational numbers.

THEOREM 1.3. *Let* ξ *be an extremal real number. Then, there exist constants* $c > 0$ *and* $t > 0$ *such that, for any rational number* $\alpha \in \mathbb{Q}$, *we have*

$$|\xi - \alpha| \geqslant cH(\alpha)^{-2}(1 + \log\, H(\alpha))^{-t}.$$

The above-mentioned Fibonacci continued fractions $\xi_{a,b}$ are an example of the more general Sturmian continued fractions studied by J.-P. Allouche, J. L. Davison, M. Queffélec and L. Q. Zamboni in [**1**]. These authors showed that Sturmian continued fractions are transcendental by showing that they admit very good approximations by quadratic real numbers. In the present situation, we get an optimal measure of approximation by such numbers.

THEOREM 1.4. *Let* ξ *be an extremal real number. There is a value of* $c > 0$ *such that the inequality*

$$|\xi - \alpha| \leqslant cH(\alpha)^{-2\gamma^2}$$

has infinitely many solutions in algebraic numbers $\alpha \in \mathbb{C}$ of degree 2, and another value of $c > 0$ such that the same inequality has no solution in algebraic numbers $\alpha \in \mathbb{C}$ of degree at most 2.

It is generally believed that, for any real number ξ that is not algebraic over $\mathbb{Q}$ of degree at most 3 and for any $\epsilon > 0$, there should exist infinitely many algebraic integers α of degree at most 3 which satisfy $|\xi - \alpha| \leqslant H(\alpha)^{-3+\epsilon}$ (see [**9**, p. 259]). A recent result of Y. Bugeaud and O. Teulié shows that this is true of almost all real numbers ξ in the sense of Lebesgue measure (see [**2**, Corollary 1 to Theorem 6]). By Schmidt's subspace theorem, this is also true of all algebraic numbers of degree greater than 3. However, we will present a criterion which suggests that this is probably not true for extremal real numbers and that, for these numbers, the correct exponent of approximation is probably γ^2 instead of 3.[†] As an application of this criterion, we will derive the following measure of approximation.

THEOREM 1.5. *Let ξ be an extremal real number. There exists a constant $c > 0$ such that, for any algebraic integer $\alpha \in \mathbb{C}$ of degree at most 3, we have*

$$|\xi - \alpha| \geqslant cH(\alpha)^{-\gamma^2-1}.$$

This does not disprove the natural conjecture stated above, but it sheds doubts on it. By extension, it also sheds doubts on the conjectures of Wirsing and Schmidt concerning approximation to a real number by algebraic numbers of bounded degree, although Davenport and Schmidt have established these conjectures in the case of approximation by quadratic irrationals in [**4**].

This paper is organized as follows. Section 5 presents a criterion for a real number to be extremal. Its proof is based on preliminary considerations from §§ 2, 3 and 4. Section 6 provides examples of extremal real numbers which include the Fibonacci continued fractions. Sections 7, 8 and 9 are devoted to measures of approximation of extremal real numbers by rational numbers, by quadratic real numbers, and by algebraic integers of degree at most 3 respectively. Section 9 also provides a criterion for approximation by algebraic integers of degree at most 3, and § 10 proves a partial converse to this criterion.

As suggested by the referee, the reader may, at a first reading, skip § 2 and part (iii) of Lemma 3.1 and go directly from Theorem 5.1 to § 7. He may also want to devise an independent proof of Proposition 6.3 showing that Fibonacci continued fractions are extremal real numbers, or look at the short note [**8**] for such a proof. However, in a more linear reading, the relevance of the algebraic tool developed in § 2 will appear clearly in Corollary 5.2, a result which the author sees as a key for further study of extremal real numbers and a motivation for the constructions of § 6.

NOTATION. Given a fixed real number ξ, we often write estimates of the form $A \ll B$ or $B \gg A$, where A and B are non-negative variable real numbers, to mean

[†] Since this paper was written the author has found a class of extremal real numbers for which this exponent of approximation is indeed γ^2; see 'Approximation to real numbers by cubic algebraic integers II', *Ann. of Math.* to appear.

$A \leqslant cB$ for some constant $c > 0$ depending only on ξ. We also write $A \sim B$ to mean $A \ll B \ll A$.

2. *A particular construction*

Let A denote a commutative ring with unit (in practice, we will take $A = \mathbb{Z}$). Identifying any point $\mathbf{x} = (x_0, x_1, x_2) \in A^3$ with the symmetric matrix

$$\mathbf{x} = \begin{pmatrix} x_0 & x_1 \\ x_1 & x_2 \end{pmatrix},$$

we can talk of its determinant

$$\det(\mathbf{x}) = x_0 x_2 - x_1^2.$$

Put

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then, given any three points $\mathbf{x}, \mathbf{y}, \mathbf{z} \in A^3$, we find, with the natural notation for their coordinates,

$$\begin{aligned} -\mathbf{x}J\mathbf{z}J\mathbf{y} &= \begin{pmatrix} x_0 & x_1 \\ x_1 & x_2 \end{pmatrix} \begin{pmatrix} z_2 & -z_1 \\ -z_1 & z_0 \end{pmatrix} \begin{pmatrix} y_0 & y_1 \\ y_1 & y_2 \end{pmatrix} \\ &= \begin{pmatrix} x_0 & x_1 \\ x_1 & x_2 \end{pmatrix} \begin{pmatrix} \begin{vmatrix} y_0 & y_1 \\ z_1 & z_2 \end{vmatrix} & \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \\ -\begin{vmatrix} y_0 & y_1 \\ z_0 & z_1 \end{vmatrix} & -\begin{vmatrix} y_1 & y_2 \\ z_0 & z_1 \end{vmatrix} \end{pmatrix} \\ &= \begin{pmatrix} \begin{vmatrix} x_0 & x_1 \\ \begin{vmatrix} y_0 & y_1 \\ z_0 & z_1 \end{vmatrix} & \begin{vmatrix} y_0 & y_1 \\ z_1 & z_2 \end{vmatrix} \end{vmatrix} & \begin{vmatrix} x_0 & x_1 \\ \begin{vmatrix} y_1 & y_2 \\ z_0 & z_1 \end{vmatrix} & \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \end{vmatrix} \\ \begin{vmatrix} x_1 & x_2 \\ \begin{vmatrix} y_0 & y_1 \\ z_0 & z_1 \end{vmatrix} & \begin{vmatrix} y_0 & y_1 \\ z_1 & z_2 \end{vmatrix} \end{vmatrix} & \begin{vmatrix} x_1 & x_2 \\ \begin{vmatrix} y_1 & y_2 \\ z_0 & z_1 \end{vmatrix} & \begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} \end{vmatrix} \end{pmatrix}. \end{aligned}$$

Moreover, if, in this matrix, we subtract from the element of row 1 and column 2 the element of row 2 and column 1, we find a polynomial in $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ which happens to be

$$\operatorname{trace}(J\mathbf{x}J\mathbf{z}J\mathbf{y}) = \det(\mathbf{x}, \mathbf{y}, \mathbf{z}). \tag{2.1}$$

So, the product $-\mathbf{x}J\mathbf{z}J\mathbf{y}$ is symmetric if and only if $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$. In this case, we denote by $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ the corresponding element of A^3:

$$[\mathbf{x}, \mathbf{y}, \mathbf{z}] = -\mathbf{x}J\mathbf{z}J\mathbf{y}. \tag{2.2}$$

LEMMA 2.1. *For any* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in A^3$ *with* $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$ *and any* $\mathbf{w} \in A^3$, *we have:*

(i) $\det[\mathbf{x}, \mathbf{y}, \mathbf{z}] = \det(\mathbf{x})\det(\mathbf{y})\det(\mathbf{z})$,
(ii) $\det(\mathbf{w}, \mathbf{y}, [\mathbf{x}, \mathbf{y}, \mathbf{z}]) = \det(\mathbf{y})\det(\mathbf{w}, \mathbf{z}, \mathbf{x})$,
(iii) $\det(\mathbf{x}, \mathbf{y}, [\mathbf{x}, \mathbf{y}, \mathbf{z}]) = 0$ *and* $[\mathbf{x}, \mathbf{y}, [\mathbf{x}, \mathbf{y}, \mathbf{z}]] = \det(\mathbf{x})\det(\mathbf{y})\mathbf{z}$.

Proof. The formula (i) follows from (2.2) since $\det(J) = 1$. To prove (ii) and (iii), we note that, for any $\mathbf{w} \in A^3$, we have

$$\mathbf{w}J\mathbf{w}J = J\mathbf{w}J\mathbf{w} = -\det(\mathbf{w})I$$

where I denotes the identity matrix. Combining this observation with (2.1), we find that

$$\begin{aligned}\det(\mathbf{w}, \mathbf{y}, [\mathbf{x}, \mathbf{y}, \mathbf{z}]) &= \mathrm{trace}(J\mathbf{w}J(-\mathbf{x}J\mathbf{z}J\mathbf{y})J\mathbf{y}) \\ &= \det(\mathbf{y})\mathrm{trace}(J\mathbf{w}J\mathbf{x}J\mathbf{z}) \\ &= \det(\mathbf{y})\det(\mathbf{w}, \mathbf{z}, \mathbf{x}).\end{aligned}$$

This proves (ii) and the first half of (iii). For the second half, we compute

$$-\mathbf{x}J(-\mathbf{x}J\mathbf{z}J\mathbf{y})J\mathbf{y} = (\mathbf{x}J\mathbf{x}J)\mathbf{z}(J\mathbf{y}J\mathbf{y}) = \det(\mathbf{x})\det(\mathbf{y})\mathbf{z}.$$ □

3. *Some estimates*

Let ξ be any real number. For each integral point $\mathbf{x} = (x_0, x_1, x_2) \in \mathbb{Z}^3$, we define

$$\|\mathbf{x}\| = \max\{|x_0|, |x_1|, |x_2|\} \quad \text{and} \quad L(\mathbf{x}) = \max\{|x_0\xi - x_1|, |x_0\xi^2 - x_2|\}.$$

For a square matrix A, we also denote by $\|A\|$ the absolute value of its determinant. In the following lemma, we recall estimates for determinants that appear in the study of Davenport and Schmidt in [**5**, § 3], and we provide estimates relative to the construction of the previous paragraph.

LEMMA 3.1. *Let* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{Z}^3$.

(i) *For any choice of integers* $r, s, t, u \in \{0, 1, 2\}$ *with* $s - r = u - t$, *we have*

$$\left\| \begin{matrix} x_r & x_s \\ y_t & y_u \end{matrix} \right\| \ll \|\mathbf{x}\|L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x}).$$

(ii) *We have*

$$|\det(\mathbf{x}, \mathbf{y}, \mathbf{z})| \ll \|\mathbf{x}\|L(\mathbf{y})L(\mathbf{z}) + \|\mathbf{y}\|L(\mathbf{x})L(\mathbf{z}) + \|\mathbf{z}\|L(\mathbf{x})L(\mathbf{y}).$$

(iii) *Put* $\mathbf{w} = [\mathbf{x}, \mathbf{x}, \mathbf{y}]$. *Then, we have*

$$\|\mathbf{w}\| \ll \|\mathbf{x}\|^2 L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x})^2 \quad \text{and} \quad L(\mathbf{w}) \ll (\|\mathbf{x}\|L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x}))L(\mathbf{x}).$$

Proof. All these estimates are based on the multilinearity of the determinant. We simply prove (iii) since (i) and (ii) follow from the computations in the proofs of Lemmas 3 and 4 of [**5**]. Write $\mathbf{w} = (w_0, w_1, w_2)$. Since

$$w_0 = \begin{vmatrix} x_0 & x_1 \\ \left|\begin{matrix} x_0 & x_1 \\ y_0 & y_1 \end{matrix}\right| & \left|\begin{matrix} x_0 & x_1 \\ y_1 & y_2 \end{matrix}\right| \end{vmatrix} = \begin{vmatrix} x_0 & x_1 - x_0\xi \\ \left|\begin{matrix} x_0 & x_1 \\ y_0 & y_1 \end{matrix}\right| & \left|\begin{matrix} x_0 & x_1 \\ y_1 - y_0\xi & y_2 - y_1\xi \end{matrix}\right| \end{vmatrix},$$

we obtain, using part (i) and noting that $L(\mathbf{x}) \ll \|\mathbf{x}\|$,

$$|w_0| \ll \|\mathbf{x}\| \left\| \begin{matrix} x_0 & x_1 \\ y_1 - y_0\xi & y_2 - y_1\xi \end{matrix} \right\| + L(\mathbf{x}) \left\| \begin{matrix} x_0 & x_1 \\ y_0 & y_1 \end{matrix} \right\| \ll \|\mathbf{x}\|^2 L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x})^2.$$

Similarly, we find, for $j = 0, 1$, that

$$|w_{j+1} - w_j\xi| = \left\| \begin{array}{cc} x_1 - x_0\xi & x_2 - x_1\xi \\ \begin{vmatrix} x_j & x_{j+1} \\ y_0 & y_1 \end{vmatrix} & \begin{vmatrix} x_j & x_{j+1} \\ y_1 & y_2 \end{vmatrix} \end{array} \right\| \ll L(\mathbf{x})(\|\mathbf{x}\|L(\mathbf{y}) + \|\mathbf{y}\|L(\mathbf{x}))$$

and the conclusion follows. □

4. *The sequence of minimal points*

Let ξ be a real number which is neither rational nor quadratic over $\mathbb{Q}$. Then, the function $L\colon \mathbb{Z}^3 \to \mathbb{R}$ defined in § 3 has the property that, for $\mathbf{x}$, $\mathbf{y} \in \mathbb{Z}^3$ with $x_0, y_0 \neq 0$, we have $L(\mathbf{x}) = L(\mathbf{y})$ if and only if $\mathbf{x} = \pm\mathbf{y}$. Accordingly, for any real number $X \geqslant 1$, the finite set of points $\mathbf{x} = (x_0, x_1, x_2)$ of $\mathbb{Z}^3$ satisfying

$$1 \leqslant x_0 \leqslant X \quad \text{and} \quad L(\mathbf{x}) = \max\{|x_0\xi - x_1|, |x_0\xi^2 - x_2|\} \leqslant 1$$

contains exactly one point for which $L(\mathbf{x})$ is minimal. Following Davenport and Schmidt, we call it the *minimal point corresponding to* X.

Clearly, if $\mathbf{x}$ is a minimal point for some real number $X \geqslant 1$, then it is a minimal point for x_0. So, we may order the minimal points according to their first coordinates. Note also that the coordinates of a minimal point are relatively prime as a set. Therefore, distinct minimal points are linearly independent over $\mathbb{Q}$ as elements of $\mathbb{Q}^3$. In particular, any two consecutive minimal points are linearly independent over $\mathbb{Q}$.

To state the next lemma, we define the height of a 2×3 matrix A, denoted $\|A\|$, to be the maximal absolute value of its minors of order 2, and we define the height of a 2-dimensional subspace V of $\mathbb{Q}^3$, denoted $H(V)$, to be the height of any 2×3 matrix whose rows form a basis of $V \cap \mathbb{Z}^3$.

LEMMA 4.1. *Let* $\mathbf{x}$ *be a minimal point, let* $\mathbf{y}$ *be the next minimal point and let* $V = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{Q}}$ *be the subspace of* $\mathbb{Q}^3$ *generated by* $\mathbf{x}$ *and* $\mathbf{y}$. *Then* $\{\mathbf{x}, \mathbf{y}\}$ *is a basis of* $V \cap \mathbb{Z}^3$ *over* $\mathbb{Z}$ *and we have* $H(V) \sim \|\mathbf{y}\|L(\mathbf{x})$.

Proof. Write $\mathbf{x} = (x_0, x_1, x_2)$ and $\mathbf{y} = (y_0, y_1, y_2)$. Arguing like Davenport and Schmidt in their proof of Lemma 2 of [**4**], we observe that, if $\{\mathbf{x}, \mathbf{y}\}$ were not a basis of $V \cap \mathbb{Z}^3$, there would exist a point $\mathbf{z} = (z_0, z_1, z_2)$ of $\mathbb{Z}^3$ with $z_0 > 0$, which could be written in the form $\mathbf{z} = r\mathbf{x} + s\mathbf{y}$ for rational numbers r and s with $\max\{|r|, |s|\} \leqslant \frac{1}{2}$. We would then get

$$z_0 \leqslant |r|x_0 + |s|y_0 < y_0 \quad \text{and} \quad L(\mathbf{z}) \leqslant |r|L(\mathbf{x}) + |s|L(\mathbf{y}) < L(\mathbf{x}),$$

in contradiction with the fact that $\mathbf{y}$ is the next minimal point after $\mathbf{x}$. Thus, $\{\mathbf{x}, \mathbf{y}\}$ is a basis of V and, using Lemma 3.1(i), we find that

$$H(V) = \left\| \begin{matrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \end{matrix} \right\| \ll \|\mathbf{y}\|L(\mathbf{x}) + \|\mathbf{x}\|L(\mathbf{y}) \ll \|\mathbf{y}\|L(\mathbf{x}).$$

To derive lower bounds for $H(V)$, put $u_i = x_i - x_0\xi^i$ and $v_i = y_i - y_0\xi^i$ for

$i = 1, 2$, and choose an index j for which $|u_j| = L(\mathbf{x})$. We obtain

$$H(V) \geqslant \left\| \begin{matrix} x_0 & x_j \\ y_0 & y_j \end{matrix} \right\| = |x_0 v_j - y_0 u_j| \geqslant y_0 |u_j| - x_0 |v_j| \geqslant (y_0 - x_0) L(\mathbf{x}).$$

Since the point $\mathbf{z} = \mathbf{y} - \mathbf{x}$ has its first coordinate $z_0 = y_0 - x_0$ satisfying $1 \leqslant z_0 < y_0$, and since $\mathbf{y}$ is the next minimal point after $\mathbf{x}$, there must exist an index i for which $|z_i - z_0 \xi^i| = |v_i - u_i| > L(\mathbf{x})$. Thus, we also find that

$$\begin{aligned} H(V) &\geqslant |x_0 v_i - y_0 u_i| \geqslant |x_0 (v_i - u_i) - (y_0 - x_0) u_i| \\ &\geqslant x_0 L(\mathbf{x}) - (y_0 - x_0) L(\mathbf{x}). \end{aligned}$$

Combining the two preceding displayed inequalities, we have

$$3H(V) \geqslant x_0 L(\mathbf{x}) + (y_0 - x_0) L(\mathbf{x}) \geqslant y_0 L(\mathbf{x}),$$

and therefore $H(V) \gg \|\mathbf{y}\| L(\mathbf{x})$. □

5. *The set of extremal numbers*

Recall that, in the introduction, we defined a real number ξ to be *extremal* if it is not rational nor quadratic irrational and if there exists a constant $c > 0$ such that the inequalities

$$1 \leqslant |x_0| \leqslant X, \quad |x_0 \xi - x_1| \leqslant c X^{-1/\gamma}, \quad |x_0 \xi^2 - x_2| \leqslant c X^{-1/\gamma}, \tag{5.1}$$

have a solution in integers x_0, x_1, x_2 for any real number $X \geqslant 1$. By virtue of the subspace theorem of Schmidt, such a real number is transcendental over $\mathbb{Q}$ (see, for example, [**9**, Chapter VI, Theorem 1B]).

In this section, we characterize the extremal real numbers by a stronger approximation property and show that the corresponding approximation triples satisfy a certain recurrence relation involving the operation $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ defined in § 2. We conclude that the set of extremal numbers is at most countable.

THEOREM 5.1. *A real number ξ is extremal if and only if there exist an increasing sequence of positive integers $(Y_k)_{k \geqslant 1}$ and a sequence of points $(\mathbf{y}_k)_{k \geqslant 1}$ of $\mathbb{Z}^3$ such that, for all $k \geqslant 1$, we have*

$$\begin{gathered} Y_{k+1} \sim Y_k^{\gamma}, \quad \|\mathbf{y}_k\| \sim Y_k, \quad L(\mathbf{y}_k) \sim Y_k^{-1}, \\ 1 \leqslant |\det(\mathbf{y}_k)| \ll 1 \quad \textit{and} \quad 1 \leqslant |\det(\mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2})| \ll 1, \end{gathered}$$

where the norm $\| \ \|$ and the map $L = L_\xi$ are as defined in § 3.

Proof. Assume first that there exist such sequences $(\mathbf{y}_k)_{k \geqslant 1}$ and $(Y_k)_{k \geqslant 1}$ for a given real number ξ. Then, ξ is not rational nor quadratic irrational because otherwise there would exist integers p, q, r, not all zero, such that $p + q\xi + r\xi^2 = 0$ and consequently, for any point $\mathbf{y}_k = (y_{k,0}, y_{k,1}, y_{k,2})$ we would have

$$|p y_{k,0} + q y_{k,1} + r y_{k,2}| = |q(y_{k,1} - y_{k,0}\xi) + r(y_{k,2} - y_{k,0}\xi^2)| \ll Y_k^{-1}.$$

Since the left-hand side of this inequality is an integer, this integer would therefore be zero for all sufficiently large k, in contradiction with the fact that the determinant of any three consecutive points of the sequence $(\mathbf{y}_k)_{k \geqslant 1}$ is non-zero. Moreover, for any sufficiently large real number X, there exists an index $k \geqslant 1$

such that $Y_k \leqslant X < Y_{k+1}$ and then the point $\mathbf{y}_k$ satisfies

$$\|\mathbf{y}_k\| \ll Y_k \ll X \quad \text{and} \quad L(\mathbf{y}_k) \ll Y_k^{-1} \ll Y_{k+1}^{-1/\gamma} \ll X^{-1/\gamma}.$$

Therefore, ξ is extremal.

Conversely, assume that ξ is an extremal real number and consider the sequence $(\mathbf{x}_i)_{i\geqslant 1}$ of minimal points of $\mathbb{Z}^3$ associated to ξ, ordered according to their first coordinate (see § 4 above). For each $i \geqslant 1$, denote by X_i the first coordinate of $\mathbf{x}_i$ and put $L_i = L(\mathbf{x}_i)$. Choose also $c > 0$ such that (5.1) has an integral solution for each $X \geqslant 1$. Then, by the definition of minimal points, we have

$$L_i \leqslant cX_{i+1}^{-1/\gamma} \quad (i \geqslant 1) \tag{5.2}$$

(see [**5**, formula (30)]). By Lemma 2 of [**5**], there is an index $i_0 \geqslant 2$ such that $\det(\mathbf{x}_i) \neq 0$ for all $i \geqslant i_0$. Moreover, [**5**, Lemma 5] shows that there are infinitely many indices $i > i_0$ such that $\mathbf{x}_{i-1}, \mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are linearly independent over $\mathbb{Q}$. Let I denote the set of all such integers i and, for each integer $k \geqslant 1$, denote by i_k the kth element of I. We claim that the sequences $(\mathbf{y}_k)_{k\geqslant 1}$ and $(Y_k)_{k\geqslant 1}$ given by

$$\mathbf{y}_k = \mathbf{x}_{i_k} \quad \text{and} \quad Y_k = X_{i_k} \quad (k \geqslant 1)$$

have the required properties.

To prove this, fix any index $k \geqslant 1$ and put $i = i_k$. The inequality (5.2) gives

$$L_{i-1} \ll X_i^{-1/\gamma} \quad \text{and} \quad L_i \ll X_{i+1}^{-1/\gamma}.$$

Then, by part (i) of Lemma 3.1 and the fact that $\det(\mathbf{x}_i)$ is a non-zero integer, we find that

$$1 \leqslant |\det(\mathbf{x}_i)| \ll X_i L_i \ll X_i X_{i+1}^{-1/\gamma} \tag{5.3}$$

and so $X_{i+1} \ll X_i^{\gamma}$. Similarly, using part (ii) of Lemma 3.1 and the fact that the integral points $\mathbf{x}_{i-1}$, $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are linearly independent over $\mathbb{Q}$, we then obtain

$$\begin{aligned} 1 \leqslant |\det(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})| &\ll X_{i+1} L_i L_{i-1} \\ &\ll X_{i+1}^{1/\gamma^2} L_{i-1} \ll X_{i+1}^{1/\gamma^2} X_i^{-1/\gamma} \ll 1. \end{aligned}$$

From these last estimates, we deduce that

$$X_{i+1} \sim X_i^{\gamma}, \quad L_{i-1} \sim X_i^{-1/\gamma} \quad \text{and} \quad L_i \sim X_{i+1}^{-1/\gamma} \sim X_i^{-1}. \tag{5.4}$$

In particular, since $\mathbf{x}_i = \mathbf{y}_k$ and $X_i = Y_k$, we obtain, from (5.3) and (5.4),

$$\|\mathbf{y}_k\| \sim Y_k, \quad L(\mathbf{y}_k) \sim Y_k^{-1} \quad \text{and} \quad 1 \leqslant |\det(\mathbf{y}_k)| \ll 1. \tag{5.5}$$

Now, let $V = \langle \mathbf{x}_i, \mathbf{x}_{i+1} \rangle_{\mathbb{Q}}$ and let $j \geqslant i+1$ be the largest integer such that V contains $\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_j$. Then we have

$$V = \langle \mathbf{x}_i, \mathbf{x}_{i+1} \rangle_{\mathbb{Q}} = \langle \mathbf{x}_{j-1}, \mathbf{x}_j \rangle_{\mathbb{Q}} \tag{5.6}$$

and, since $\mathbf{x}_{j+1} \notin V$, we deduce that the points $\mathbf{x}_{j-1}$, $\mathbf{x}_j$ and $\mathbf{x}_{j+1}$ are linearly independent over $\mathbb{Q}$. Moreover, since $\mathbf{x}_i, \ldots, \mathbf{x}_j \in V$, the index j is the smallest integer greater than i with the latter property. Therefore, we have $j = i_{k+1}$. Consequently the estimates (5.4) also apply with i replaced by j. Lemma 4.1

then gives

$$H(\langle \mathbf{x}_i, \mathbf{x}_{i+1}\rangle_{\mathbb{Q}}) \sim X_{i+1}L_i \sim X_i^{1/\gamma}$$

and

$$H(\langle \mathbf{x}_{j-1}, \mathbf{x}_j\rangle_{\mathbb{Q}}) \sim X_jL_{j-1} \sim X_j^{1/\gamma^2}.$$

By virtue of (5.6), this implies that $X_j \sim X_i^{\gamma}$ and, since $X_i = Y_k$ and $X_j = Y_{k+1}$, we conclude that

$$Y_{k+1} \sim Y_k^{\gamma}. \tag{5.7}$$

The preceding argument also shows that $\langle \mathbf{y}_k, \mathbf{y}_{k+1}\rangle_{\mathbb{Q}}$ contains $\mathbf{x}_{i+1}$ and $\mathbf{x}_{j-1}$. Therefore, $\langle \mathbf{y}_{k+1}, \mathbf{y}_{k+2}\rangle_{\mathbb{Q}}$ contains $\mathbf{x}_{j+1}$ and so $\langle \mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2}\rangle_{\mathbb{Q}}$ contains the three linearly independent points $\mathbf{x}_{j-1}$, $\mathbf{x}_j$ and $\mathbf{x}_{j+1}$. This implies that $\mathbf{y}_k$, $\mathbf{y}_{k+1}$ and $\mathbf{y}_{k+2}$ are linearly independent over $\mathbb{Q}$ and, combining Lemma 3.1(ii) with (5.5) and (5.7), we find that

$$1 \leqslant |\det(\mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2})| \ll \|\mathbf{y}_{k+2}\|L(\mathbf{y}_k)L(\mathbf{y}_{k+1}) \ll Y_{k+2}Y_k^{-1}Y_{k+1}^{-1} \ll 1.$$

This estimate, together with (5.5) and (5.7), proves that the sequences $(\mathbf{y}_k)_{k\geqslant 1}$ and $(Y_k)_{k\geqslant 1}$ have all the required properties. □

COROLLARY 5.2. *Let ξ be an extremal real number and let $(\mathbf{y}_k)_{k\geqslant 1}$ and $(Y_k)_{k\geqslant 1}$ be as in the statement of Theorem* 5.1. *Then, for any sufficiently large integer $k \geqslant 3$, the point $\mathbf{y}_{k+1}$ is a non-zero rational multiple of $[\mathbf{y}_k, \mathbf{y}_k, \mathbf{y}_{k-2}]$.*

Proof. Let $k \geqslant 4$ be an integer and let $\mathbf{w} = [\mathbf{y}_k, \mathbf{y}_k, \mathbf{y}_{k+1}]$. By Lemma 2.1(i), we have

$$\det(\mathbf{w}) = \det(\mathbf{y}_k)^2\det(\mathbf{y}_{k+1}).$$

Thus $\det(\mathbf{w}) \neq 0$ and consequently, $\mathbf{w}$ is a non-zero point of $\mathbb{Z}^3$. Using Lemma 3.1(iii), we find that

$$\|\mathbf{w}\| \ll Y_k^2Y_{k+1}^{-1} \sim Y_{k-2} \quad \text{and} \quad L(\mathbf{w}) \ll Y_{k+1}Y_k^{-2} \sim Y_{k-2}^{-1}.$$

Then, applying Lemma 3.1(ii), we obtain

$$|\det(\mathbf{w}, \mathbf{y}_{k-3}, \mathbf{y}_{k-2})| \ll Y_{k-2}Y_{k-2}^{-1}Y_{k-3}^{-1} \sim Y_{k-3}^{-1},$$
$$|\det(\mathbf{w}, \mathbf{y}_{k-2}, \mathbf{y}_{k-1})| \ll Y_{k-1}Y_{k-2}^{-2} \sim Y_{k-3}^{-1/\gamma}.$$

But the above two determinants are integers. So, if k is sufficiently large, they must be zero. Since $\mathbf{y}_{k-3}$, $\mathbf{y}_{k-2}$ and $\mathbf{y}_{k-1}$ are linearly independent over $\mathbb{Q}$, this then forces $\mathbf{w}$ to be a rational multiple of $\mathbf{y}_{k-2}$. Since, by Lemma 2.1(iii), we have

$$[\mathbf{y}_k, \mathbf{y}_k, \mathbf{w}] = \det(\mathbf{y}_k)^2\mathbf{y}_{k+1},$$

we conclude that $\mathbf{y}_{k+1}$ is a rational multiple of $[\mathbf{y}_k, \mathbf{y}_k, \mathbf{y}_{k-2}]$. □

As we will see in the next section, this corollary provides a way to construct extremal numbers. It also provides the following formula.

COROLLARY 5.3. *In the notation of Corollary* 5.2, *we have*

$$\det(\mathbf{y}_{k-2},\mathbf{y}_{k-1},\mathbf{y}_k)\mathbf{y}_{k+1} = \det(\mathbf{y}_{k-2},\mathbf{y}_{k-1},\mathbf{y}_{k+1})\mathbf{y}_k + \det(\mathbf{y}_{k-1},\mathbf{y}_k,\mathbf{y}_{k+1})\mathbf{y}_{k-2}$$

for any sufficiently large value of k.

Proof. By Lemma 2.1(ii), we have $\det(\mathbf{y}_{k-2},\mathbf{y}_k,[\mathbf{y}_k,\mathbf{y}_k,\mathbf{y}_{k-2}]) = 0$ and this, by Corollary 5.2, implies that the points $\mathbf{y}_{k-2}$, $\mathbf{y}_k$ and $\mathbf{y}_{k+1}$ are linearly dependent over $\mathbb{Q}$. Writing $\mathbf{y}_{k+1} = a\mathbf{y}_k + b\mathbf{y}_{k-2}$ with unknown rational coefficients a and b and using the multilinearity of the determinant, we get

$$\det(\mathbf{y}_{k-2},\mathbf{y}_{k-1},\mathbf{y}_{k+1}) = a\det(\mathbf{y}_{k-2},\mathbf{y}_{k-1},\mathbf{y}_k)$$

and

$$\det(\mathbf{y}_{k-1},\mathbf{y}_k,\mathbf{y}_{k+1}) = b\det(\mathbf{y}_{k-1},\mathbf{y}_k,\mathbf{y}_{k-2})$$

and the formula follows. □

COROLLARY 5.4. *The set* $\mathcal{E}$ *of extremal real numbers is at most countable.*

Proof. We prove this by constructing an injective map $\varphi\colon \mathcal{E} \to (\mathbb{Z}^3)^3$ as follows. For each extremal real number $\xi \in \mathcal{E}$, we choose a corresponding sequence $(\mathbf{y}_k)_{k\geqslant 1}$ as in Theorem 5.1. Then, in accordance with Corollary 5.3, we choose an index $i \geqslant 3$ such that, for each $k \geqslant i$, the point $\mathbf{y}_{k+1}$ is a non-zero rational multiple of $[\mathbf{y}_k,\mathbf{y}_k,\mathbf{y}_{k-2}]$, and we put $\varphi(\xi) = (\mathbf{y}_{i-2},\mathbf{y}_{i-1},\mathbf{y}_i)$. Since the knowledge of $\mathbf{y}_{i-2}$, $\mathbf{y}_{i-1}$ and $\mathbf{y}_i$ determines all points $\mathbf{y}_k$ with $k \geqslant i-2$ up to a non-zero rational multiple, it determines uniquely the corresponding ratios $y_{k,1}/y_{k,0}$ and also their limit $\xi = \lim_{k\to\infty} y_{k,1}/y_{k,0}$. Thus, φ is injective. □

In order to complete the proof of Theorem 1.1, it remains to show that the set of extremal numbers is infinite. This is the object of the next section.

6. *Construction of extremal numbers*

As in § 2, we identify each point (y_0,y_1,y_2) of $\mathbb{Z}^3$ with the corresponding symmetric matrix with integral coefficients $\begin{pmatrix} y_0 & y_1 \\ y_1 & y_2 \end{pmatrix}$. Before stating the main result of this section, we first prove a lemma.

LEMMA 6.1. *Let* A *and* B *be non-commuting symmetric matrices in* $\mathrm{GL}_2(\mathbb{Z})$. *Consider the sequence* $(\mathbf{y}_k)_{k\geqslant -1}$ *constructed recursively by putting*

$$\begin{aligned} &\mathbf{y}_{-1} = B^{-1}, \quad \mathbf{y}_0 = I, \quad \mathbf{y}_1 = A, \\ &\mathbf{y}_k = [\mathbf{y}_{k-1},\mathbf{y}_{k-1},\mathbf{y}_{k-3}] \quad \text{for } k \geqslant 2, \end{aligned} \tag{6.1}$$

where I *denotes the identity* 2×2 *matrix. Then, for each integer* $k \geqslant -1$, *we have*

$$|\det(\mathbf{y}_k)| = 1 \quad \text{and} \quad |\det(\mathbf{y}_k,\mathbf{y}_{k+1},\mathbf{y}_{k+2})| = |\mathrm{trace}(JAB)| \neq 0 \tag{6.2}$$

where J *is as defined in* § 2. *Moreover, for* $k \geqslant 1$, *we have the alternative*

recurrence relation

$$\mathbf{y}_k = \pm \mathbf{y}_{k-1} S \mathbf{y}_{k-2} \quad \text{where } S = \begin{cases} AB & \text{if } k \text{ is odd,} \\ BA & \text{if } k \text{ is even.} \end{cases} \tag{6.3}$$

Proof. For $k \geqslant 2$, the recurrence relation in (6.1) implies, using Lemma 2.1(i) and (ii),

$$\det(\mathbf{y}_k) = \det(\mathbf{y}_{k-1})^2 \det(\mathbf{y}_{k-3}),$$
$$\det(\mathbf{y}_{k-2}, \mathbf{y}_{k-1}, \mathbf{y}_k) = -\det(\mathbf{y}_{k-1}) \det(\mathbf{y}_{k-3}, \mathbf{y}_{k-2}, \mathbf{y}_{k-1}).$$

Since A and B belong to $\mathrm{GL}_2(\mathbb{Z})$, we also have $\det(\mathbf{y}_k) = \pm 1$ for $k = -1, 0, 1$, and therefore we deduce, by recurrence, that, for all $k \geqslant -1$, we have

$$|\det(\mathbf{y}_k)| = 1 \quad \text{and} \quad |\det(\mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2})| = |\det(\mathbf{y}_{-1}, \mathbf{y}_0, \mathbf{y}_1)|.$$

Moreover, using (2.1), we find that

$$\det(\mathbf{y}_{-1}, \mathbf{y}_0, \mathbf{y}_1) = \det(B) \mathrm{trace}(JAB) \neq 0$$

since $AB \neq BA$. This proves (6.2).

Similarly, a short computation shows that (6.3) holds for $k = 1$ and, assuming that it holds for some integer $k \geqslant 1$, we find, using (2.2), that

$$\mathbf{y}_{k+1} = -\mathbf{y}_k J \mathbf{y}_{k-2} J \mathbf{y}_k = \pm \mathbf{y}_{k-1} S \mathbf{y}_{k-2} J \mathbf{y}_{k-2} J \mathbf{y}_k = \pm \mathbf{y}_{k-1} S \mathbf{y}_k$$

which, by taking transposes, implies that $\mathbf{y}_{k+1} = \pm \mathbf{y}_k{}^t S \mathbf{y}_{k-1}$. Thus (6.3) holds for all $k \geqslant 1$. □

THEOREM 6.2. *Let A, B and the sequence $(\mathbf{y}_k)_{k \geqslant -1}$ be as in Lemma 6.1. Assume that all coefficients of A are non-negative and that all those of AB are positive. Then, upon writing $\mathbf{y}_k = (y_{k,0}, y_{k,1}, y_{k,2})$, we find that $y_{k,0}$ is a non-zero integer for all $k \geqslant 2$ and that the limit*

$$\xi = \lim_{k \to \infty} y_{k,1}/y_{k,0}$$

exists and is an extremal real number. Moreover, letting $Y_k = \|\mathbf{y}_k\|$ for all k, we see that the sequences $(\mathbf{y}_k)_{k \geqslant 3}$ and $(Y_k)_{k \geqslant 3}$ share, with respect to ξ, all the properties stated in Theorem 5.1.

Proof. By virtue of (6.2), the sequence $(\mathbf{y}_k)_{k \geqslant -1}$ satisfies the last two conditions of Theorem 5.1. Since $\mathbf{y}_0$ and $\mathbf{y}_1$ have non-negative coefficients, since their rows and columns are all non-zero, and since AB and its transpose BA have positive coefficients, the recurrence relation (6.3) implies that, for $k \geqslant 2$, the point $\mathbf{y}_k$ has non-zero coefficients, all of the same sign. Thus, the integer $Y_k = \|\mathbf{y}_k\|$ is positive for $k \geqslant 2$ and there exists a constant $c_1 > 1$ such that, for all $k \geqslant 4$, we have

$$Y_{k-2} Y_{k-1} < Y_k \leqslant c_1 Y_{k-2} Y_{k-1}. \tag{6.4}$$

In particular, the sequence $(Y_k)_{k \geqslant 3}$ is unbounded and monotone increasing. Similarly, we find that $\lim_{k \to \infty} y_{k,0} = \infty$.

For each $k \geqslant 2$, define I_k to be the interval of $\mathbb{R}$ with end points $y_{k,1}/y_{k,0}$ and $y_{k,2}/y_{k,1}$. For $k \geqslant 3$, the relation (6.3) shows that there are positive integers r, s, t, u

depending on k such that

$$\begin{pmatrix} y_{k,0} & y_{k,1} \\ y_{k,1} & y_{k,2} \end{pmatrix} = \pm \begin{pmatrix} y_{k-1,0} & y_{k-1,1} \\ y_{k-1,1} & y_{k-1,2} \end{pmatrix} \begin{pmatrix} r & s \\ t & u \end{pmatrix}$$

and consequently we find that

$$\frac{y_{k,1}}{y_{k,0}} = \frac{ry_{k-1,1} + ty_{k-1,2}}{ry_{k-1,0} + ty_{k-1,1}} \in I_{k-1} \quad \text{and} \quad \frac{y_{k,2}}{y_{k,1}} = \frac{sy_{k-1,1} + uy_{k-1,2}}{sy_{k-1,0} + uy_{k-1,1}} \in I_{k-1}.$$

Thus, the intervals I_k with $k \geqslant 2$ form a non-increasing sequence $I_2 \supseteq I_3 \supseteq I_4 \supseteq \dots$ and, since the length of I_k is $1/(y_{k,0}y_{k,1})$ tending to zero as k tends to infinity, their intersection is reduced to one point ξ with $\xi > 0$. The ratios $y_{k,1}/y_{k,0}$ and $y_{k,2}/y_{k,1}$ therefore converge to that number ξ and, since $\xi \neq 0$, we get

$$|y_{k,0}| \sim |y_{k,1}| \sim |y_{k,2}| \sim Y_k.$$

In particular, the length of I_k satisfies $|I_k| = (y_{k,0}y_{k,1})^{-1} \sim Y_k^{-2}$ and so we find that

$$L(\mathbf{y}_k) \sim \max_{j=0,1} |y_{k,j}\xi - y_{k,j+1}| \sim Y_k \max_{j=0,1} \left| \xi - \frac{y_{k,j+1}}{y_{k,j}} \right| \sim Y_k |I_k| \sim Y_k^{-1}.$$

To conclude from Theorem 5.1 that ξ is extremal, it remains only to show that $Y_k \sim Y_{k-1}^{\gamma}$. To this end, we observe that, for $k \geqslant 4$, the real number $q_k = Y_k Y_{k-1}^{-\gamma}$ satisfies, by (6.4),

$$q_{k-1}^{-1/\gamma} \leqslant q_k \leqslant c_1 q_{k-1}^{-1/\gamma}.$$

By recurrence on k, this implies that, for all $k \geqslant 3$, we have $c_2^{-1/\gamma} \leqslant q_k \leqslant c_2$ for the constant $c_2 = \max\{c_1^{\gamma}, q_3, q_3^{-\gamma}\}$ and so $c_2^{-1/\gamma} Y_{k-1}^{\gamma} \leqslant Y_k \leqslant c_2 Y_{k-1}^{\gamma}$ as required. □

Let $E = \{a, b\}$ be an alphabet of two letters and let E^* denote the monoid of words on E for the product given by concatenation. The *Fibonacci sequence* in E^* is the sequence of words $(w_i)_{i \geqslant 0}$ defined recursively by

$$w_0 = b, \quad w_1 = a \quad \text{and} \quad w_i = w_{i-1}w_{i-2} \ \ (i \geqslant 2)$$

(see [**7**, Example 1.3.6]). Since, for every $i \geqslant 1$, the word w_i is a prefix of w_{i+1}, this sequence converges to an infinite word $w = abaabab\dots$ called the *Fibonacci word* on $\{a, b\}$. We now turn to Fibonacci continued fractions.

COROLLARY 6.3. *Let a and b be distinct positive integers and let*

$$\xi_{a,b} = [0, a, b, a, a, b, a, \dots] = 1/(a + 1/(b + \dots))$$

be the real number whose sequence of partial quotients starts with 0 followed by the elements of the Fibonacci word on $\{a, b\}$. Then $\xi_{a,b}$ is an extremal real number. It is the special case of the construction of Theorem 6.2 corresponding to the choice of

$$A = \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b & 1 \\ 1 & 0 \end{pmatrix}.$$

Note that, if ξ is an extremal real number, then $1/\xi$ is also extremal. So, the corollary implies that $1/\xi_{a,b} = [a, b, a, a, b, \dots]$ as well is extremal. More generally,

one finds that, if ξ is extremal, then the ratio $(r\xi+s)/(t\xi+u)$ is also extremal for any choice of rational numbers r, s, t, u with $ru-st\neq 0$.

Proof. Recall that, if a real number ξ has an infinite continued fraction expansion $\xi=[a_0,a_1,a_2,\ldots]$ and if, for each $i\geqslant 0$, we write its convergent $[a_0,a_1,\ldots,a_i]$ in the form p_i/q_i with relatively prime integers p_i and $q_i>0$, then the sequences $(p_i)_{i\geqslant 0}$ and $(q_i)_{i\geqslant 0}$ satisfy recurrence relations which, in matrix form, can be written for $i\geqslant 2$ as

$$\begin{pmatrix} q_i & q_{i-1} \\ p_i & p_{i-1} \end{pmatrix} = \begin{pmatrix} q_{i-1} & q_{i-2} \\ p_{i-1} & p_{i-2} \end{pmatrix} \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}$$

(see, for example, [**9** Chapter I]). When $a_0=0$, this gives

$$\begin{pmatrix} q_i & q_{i-1} \\ p_i & p_{i-1} \end{pmatrix} = \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_2 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}.$$

Now, consider the morphism of monoids $\varphi\colon E^*\to \mathrm{GL}_2(\mathbb{Z})$ determined by the conditions

$$\varphi(a)=A \quad \text{and} \quad \varphi(b)=B.$$

For each integer $k\geqslant 0$, let m_k denote the word w_{k+2} minus its last two letters and let $\mathbf{x}_k=\varphi(m_k)$. Then, we have $m_0=1$ (the empty word), $m_1=a$ and, for $k\geqslant 2$,

$$m_k = m_{k-1}sm_{k-2} \quad \text{where } s=\begin{cases} ab & \text{if } k \text{ is odd,} \\ ba & \text{if } k \text{ is even.} \end{cases}$$

In terms of the matrices $\mathbf{x}_k$, this implies that $\mathbf{x}_0=I, \mathbf{x}_1=A$ and, for $k\geqslant 2$,

$$\mathbf{x}_k = \mathbf{x}_{k-1}S\mathbf{x}_{k-2} \quad \text{where } S=\begin{cases} AB & \text{if } k \text{ is odd,} \\ BA & \text{if } k \text{ is even.} \end{cases}$$

Therefore, if $(\mathbf{y}_k)_{k\geqslant -1}$ denotes the sequence furnished by Lemma 6.1 for the above choice of A and B, we deduce from (6.3) that $\mathbf{y}_k=\pm\mathbf{x}_k$ for any $k\geqslant 0$. In particular, $\mathbf{x}_k$ is a symmetric matrix and, with the usual notation for coordinates, we have

$$\frac{y_{k,1}}{y_{k,0}} = \frac{x_{k,1}}{x_{k,0}}$$

for all $k\geqslant 1$. On one hand, Theorem 5.1 tells us that the above ratio converges to an extremal real number ξ as k tends to infinity. On the other hand, since m_k is a prefix of the Fibonacci word w and since $\mathbf{x}_k=\varphi(m_k)$, the observation made at the beginning of the proof shows that the same ratio is the convergent of $\xi_{a,b}$ whose partial quotients are 0 followed by the elements of m_k. Therefore $\xi_{a,b}=\xi$ is extremal. □

7. *Approximation by rational numbers*

In this section, we prove Theorem 1.3, as a consequence of the following result.

PROPOSITION 7.1. *Let ξ be an extremal real number. There exists a positive constant s such that, for any sufficiently large real number X, the first minimum*

of the convex body

$$\mathcal{C}(X) = \{(x_0, x_1) \in \mathbb{R}^2 : |x_0| \leqslant X \text{ and } |x_0\xi - x_1| \leqslant X^{-1}\}$$

is bounded below by $(\log X)^{-s}$.

Proof. Let $(Y_k)_{k\geqslant 1}$ and $(\mathbf{y}_k)_{k\geqslant 1}$ be the sequences given by Theorem 5.1. Let X be a real number with $X \geqslant Y_2$, let $\lambda = \lambda(X)$ be the first minimum of $\mathcal{C}(X)$, and let (x_0, x_1) be a point of $\mathbb{Z}^2$ which realizes this minimum, so that we have $|x_0| \leqslant \lambda X$ and $|x_0\xi - x_1| \leqslant \lambda X^{-1}$. We choose an index $k \geqslant 2$ such that $Y_k \leqslant X \leqslant Y_{k+1}$ and consider the point (z_0, z_1) of $\mathbb{Z}^2$ given by

$$(z_0, z_1) = (x_0, x_1) J\mathbf{y}_{k+1} J\mathbf{y}_{k-1}$$

in the notation of § 2. Writing $\mathbf{y}_{k-1} = (y_0^*, y_1^*, y_2^*)$ and $\mathbf{y}_{k+1} = (y_0', y_1', y_2')$, and using the estimates of Theorem 5.1 together with Lemma 3.1(i), we find that

$$\begin{aligned} |z_0| &= \left\| \begin{vmatrix} x_0 & x_1 \\ y_0^* & y_1^* \\ y_0' & y_1' \end{vmatrix} \quad \begin{vmatrix} y_0^* & y_1^* \\ y_1' & y_2' \end{vmatrix} \right\| \\ &= \left\| \begin{vmatrix} x_0 & x_0\xi - x_1 \\ y_0^* & y_1^* \\ y_0' & y_1' \end{vmatrix} \quad \begin{vmatrix} y_0^* & y_1^* \\ y_0'\xi - y_1' & y_1'\xi - y_2' \end{vmatrix} \right\| \\ &\ll \frac{|x_0|}{Y_k} + Y_k|x_0\xi - x_1| \ll \frac{\lambda X}{Y_k} \end{aligned}$$

and

$$\begin{aligned} |z_0\xi - z_1| &= \left\| \begin{vmatrix} x_0 & x_1 \\ y_0^*\xi - y_1^* & y_1^*\xi - y_2^* \\ y_0' & y_1' \end{vmatrix} \quad \begin{vmatrix} y_0^*\xi - y_1^* & y_1^*\xi - y_2^* \\ y_1' & y_2' \end{vmatrix} \right\| \\ &= \left\| \begin{vmatrix} x_0 & x_0\xi - x_1 \\ y_0^*\xi - y_1^* & y_1^*\xi - y_2^* \\ y_0' & y_1' \end{vmatrix} \quad \begin{vmatrix} y_0^*\xi - y_1^* & y_1^*\xi - y_2^* \\ y_0'\xi - y_1' & y_1'\xi - y_2' \end{vmatrix} \right\| \\ &\ll \frac{|x_0|}{Y_{k-1}Y_{k+1}} + Y_k|x_0\xi - x_1| \ll \frac{\lambda Y_k}{X}. \end{aligned}$$

Thus, if we put $Z = 2X/Y_k$, these relations imply that the first minimum $\lambda(Z)$ of $\mathcal{C}(Z)$ satisfies $\lambda(Z) \leqslant c\lambda(X)$ for some constant $c > 0$ which is independent of k and X. We also note that $Z \ll X^{1/\gamma^2}$ and thus $Z \leqslant X^{1/2}$ if X is sufficiently large, say $X \geqslant X_0 \geqslant Y_2$. Then, by choosing $s > 0$ with $2^{s-1} \geqslant c$, we get, for $X \geqslant X_0$,

$$\lambda(X)(\log X)^s \geqslant 2\lambda(Z)(\log Z)^s$$

where Z is a real number in the interval $2 \leqslant Z \leqslant X^{1/2}$. Since $\lambda(X)(\log X)^s$ is bounded below by some positive constant for $2 \leqslant X \leqslant X_0^{1/2}$, this shows that $\lambda(X)(\log X)^s$ tends to infinity with X. □

Proof of Theorem 1.3. Let s be as in the proposition and assume that a rational number $\alpha = p/q$ with denominator $q \geqslant 3$ satisfies $|\xi - p/q| \leqslant q^{-2}(\log q)^{-2s-2}$. Then, by putting $X = q(\log q)^{s+1}$ and $\lambda = (\log q)^{-s-1}$, we get $|q| \leqslant \lambda X$ and $|q\xi - p| \leqslant \lambda/X$ so that the first minimum of $\mathcal{C}(X)$ is at most λ. By the proposition, this implies

that $\lambda \geqslant (\log X)^{-s}$ if q is sufficiently large. This in turn forces an upper bound on q and therefore on $H(p/q)$. Thus Theorem 1.3 holds with $t = 2s + 2$. □

8. *Approximation by quadratic real numbers*

In this section, we prove Theorems 1.2 and 1.4 as consequences of the following result.

PROPOSITION 8.1. *Let ξ be an extremal real number and let the sequences $(Y_k)_{k\geqslant 1}$ and $(\mathbf{y}_k)_{k\geqslant 1}$ be as in Theorem 5.1. For each $k \geqslant 1$, we define*

$$Q_k(T) = \begin{vmatrix} 1 & T & T^2 \\ y_{k,0} & y_{k,1} & y_{k,2} \\ y_{k+1,0} & y_{k+1,1} & y_{k+1,2} \end{vmatrix}.$$

Then, Q_k is a polynomial with integral coefficients which, for all sufficiently large values of k, has degree 2 and satisfies

$$H(Q_k) \sim |Q_k'(\xi)| \sim Y_{k-1} \quad \text{and} \quad |Q_k(\xi)| \sim Y_{k+2}^{-1}.$$

Proof. By Lemma 3.1(i), we have

$$H(Q_k) = \left\| \begin{matrix} y_{k,0} & y_{k,1} & y_{k,2} \\ y_{k+1,0} & y_{k+1,1} & y_{k+1,2} \end{matrix} \right\| \ll Y_{k+1}L(\mathbf{y}_k) + Y_kL(\mathbf{y}_{k+1}) \ll Y_{k-1}.$$

All the remaining estimates assume that k is sufficiently large. Since

$$Q_k'' = 2\begin{vmatrix} y_{k,0} & y_{k,1} - y_{k,0}\xi \\ y_{k+1,0} & y_{k+1,1} - y_{k+1,0}\xi \end{vmatrix} = 2y_{k+1,0}(y_{k,0}\xi - y_{k,1}) + \mathcal{O}(Y_{k-1}^{-1}),$$

we get $Q_k'' \neq 0$ for $k \gg 1$ using the lower bound on $|y_{k,0}\xi - y_{k,1}|$ provided by Theorem 1.3. Similarly, we have

$$Q_k'(\xi) = -\begin{vmatrix} y_{k,0} & y_{k,2} - 2y_{k,1}\xi + y_{k,0}\xi^2 \\ y_{k+1,0} & y_{k+1,2} - 2y_{k+1,1}\xi + y_{k+1,0}\xi^2 \end{vmatrix}. \tag{8.1}$$

Note that, for any $j \geqslant 1$,

$$y_{j,0}(y_{j,2} - 2y_{j,1}\xi + y_{j,0}\xi^2) = \det(\mathbf{y}_j) + (y_{j,1} - y_{j,0}\xi)^2 = \det(\mathbf{y}_j) + \mathcal{O}(Y_j^{-2}),$$

and therefore

$$|y_{j,2} - 2y_{j,1}\xi + y_{j,0}\xi^2| \sim Y_j^{-1}.$$

Using this to estimate (8.1), we find that

$$|Q_k'(\xi)| = |y_{k+1,0}|\,|y_{k,2} - 2y_{k,1}\xi + y_{k,0}\xi^2| + \mathcal{O}(Y_{k-1}^{-1}) \sim Y_{k-1}$$

and, since $H(Q_k) \ll Y_{k-1}$, we get $H(Q_k) \sim Y_{k-1}$. Finally, write

$$y_{k+2,0}(1, \xi, \xi^2) = \mathbf{y}_{k+2} + \mathbf{z}$$

where $\mathbf{z} \in \mathbb{R}^3$ has norm $\|\mathbf{z}\| \ll Y_{k+2}^{-1}$. By multilinearity of the determinant, this gives

$$y_{k+2,0}Q_k(\xi) = \det(\mathbf{y}_{k+2}, \mathbf{y}_k, \mathbf{y}_{k+1}) + \det(\mathbf{z}, \mathbf{y}_k, \mathbf{y}_{k+1}).$$

Since

$$|\det(\mathbf{z}, \mathbf{y}_k, \mathbf{y}_{k+1})| \leqslant 3\|\mathbf{z}\|H(Q_k) \ll Y_{k-1}/Y_{k+2},$$

we conclude that

$$|Q_k(\xi)| = |y_{k+2,0}|^{-1}(|\det(\mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2})| + \mathcal{O}(Y_{k-1}/Y_{k+2})) \sim Y_{k+2}^{-1}. \qquad \square$$

Proof of Theorem 1.2. For any real number X with $X \geqslant H(Q_1)$, we choose an index $k \geqslant 1$ such that $H(Q_k) \leqslant X \leqslant H(Q_{k+1})$. Then upon writing $Q_k(T) = x_0T^2 + x_1T + x_2$, we find that (x_0, x_1, x_2) is a non-zero integral point with $|x_0| \leqslant X$, $|x_1| \leqslant X$ and

$$|x_0\xi^2 + x_1\xi + x_2| = |Q_k(\xi)| \sim Y_{k+2}^{-1} \sim H(Q_{k+1})^{-\gamma^2} \ll X^{-\gamma^2}. \qquad \square$$

As for Theorem 1.4, it follows immediately from the following more precise result.

THEOREM 8.2. *Let ξ be an extremal real number. Then there exist an integer $k_0 \geqslant 1$ and positive constants c_1 and c_2 such that, for any $k \geqslant k_0$, the polynomial Q_k of Lemma* 8.1 *is irreducible over $\mathbb{Q}$ of degree* 2 *and admits exactly one root α_k with*

$$c_1H(\alpha_k)^{-2\gamma^2} \leqslant |\xi - \alpha_k| \leqslant c_2H(\alpha_k)^{-2\gamma^2}. \tag{8.2}$$

There also exists a constant $c_3 > 0$ such that, for any algebraic number $\alpha \in \mathbb{C}$ of degree at most 2 *over $\mathbb{Q}$, distinct from all α_k with $k \geqslant k_0$, we have*

$$|\xi - \alpha| \geqslant c_3H(\alpha)^{-4}.$$

Proof. By Proposition 8.1, the polynomial Q_k has degree 2 for $k \gg 1$. For these values of k, we may factor it as $Q_k(T) = a_k(T - \alpha_k)(T - \beta_k)$ with the roots α_k and β_k ordered so that $|\xi - \alpha_k| \leqslant |\xi - \beta_k|$. Then the logarithmic derivative of Q_k satisfies

$$\frac{1}{|\xi - \alpha_k|} - \frac{1}{|\xi - \beta_k|} \leqslant \frac{|Q_k'(\xi)|}{|Q_k(\xi)|} = \left|\frac{1}{\xi - \alpha_k} + \frac{1}{\xi - \beta_k}\right| \leqslant \frac{2}{|\xi - \alpha_k|}. \tag{8.3}$$

Using the estimates of the proposition, this gives

$$|\xi - \alpha_k| \leqslant 2\frac{|Q_k(\xi)|}{|Q_k'(\xi)|} \sim H(Q_k)^{-2\gamma^2} \ll H(\alpha_k)^{-2\gamma^2}. \tag{8.4}$$

Since $H(Q_k)$ tends to infinity with k, we deduce that α_k tends to ξ as k tends to infinity and, by comparison with Theorem 1.3, that α_k is irrational for all sufficiently large values of k. Thus, for $k \gg 1$, the polynomial Q_k is irreducible over $\mathbb{Q}$. Its discriminant $a_k^2(\alpha_k - \beta_k)^2$ is then a non-zero integer, and so we find that

$$|\alpha_k - \beta_k| \geqslant |a_k|^{-1} \geqslant H(Q_k)^{-1}.$$

In particular, we have $|\alpha_k - \beta_k| \geqslant 3|\xi - \alpha_k|$ for any sufficiently large value of k,

and (8.3) becomes

$$\frac{1}{2|\xi-\alpha_k|} \leqslant \frac{|Q_k'(\xi)|}{|Q_k(\xi)|} \leqslant \frac{2}{|\xi-\alpha_k|}.$$

On the other hand, the greatest common divisor of the coefficients of Q_k divides the non-zero integer $\det(\mathbf{y}_{k-1}, \mathbf{y}_k, \mathbf{y}_{k+1})$ and, since the absolute value of this integer is bounded above by a constant that is independent of k, we also have $H(\alpha_k) \sim H(Q_k)$ for $k \gg 1$. Then, going back to (8.4), we get refined estimates of the form (8.2) valid, say, for all $k \geqslant k_0$, where c_1 and c_2 are positive constants and $k_0 \geqslant 2$ is a fixed integer. This proves the first part of the theorem.

Now, let $\alpha \in \mathbb{C}$ be an algebraic number of degree at most 2 with $\alpha \neq \alpha_k$ for any index $k \geqslant k_0$. For these values of k, Liouville's inequality gives

$$|\alpha-\alpha_k| \geqslant c_3 H(\alpha)^{-2} H(\alpha_k)^{-2}$$

where $c_3 > 0$ is an absolute constant (see, for example, [**6**, Lemma 3]). Choose k to be the smallest integer not less than k_0 satisfying

$$H(\alpha) \leqslant \left(\frac{c_3}{2c_2}\right)^{1/2} H(\alpha_k)^{\gamma}.$$

This choice of k ensures that

$$|\xi-\alpha_k| \leqslant c_2 H(\alpha_k)^{-2\gamma^2} \leqslant \tfrac{1}{2} c_3 H(\alpha)^{-2} H(\alpha_k)^{-2} \leqslant \tfrac{1}{2}|\alpha-\alpha_k|,$$

and therefore that

$$|\xi-\alpha| \geqslant \tfrac{1}{2}|\alpha-\alpha_k| \gg H(\alpha)^{-2} H(\alpha_k)^{-2}.$$

If $k > k_0$, we also have $H(\alpha) \gg H(\alpha_{k-1})^{\gamma} \gg H(\alpha_k)$, while if $k = k_0$, we have $H(\alpha) \geqslant 1 \gg H(\alpha_k)$. So, the above inequality leads to $|\xi-\alpha| \gg H(\alpha)^{-4}$. □

9. *Approximation by cubic algebraic integers*

For a real number x, let $\{x\} = \min\{|x-p| : p \in \mathbb{Z}\}$ denote the distance from x to a nearest integer. In this section, we fix an extremal real number ξ and corresponding sequences $(Y_k)_{k\geqslant 1}$ and $(\mathbf{y}_k)_{k\geqslant 1}$ satisfying the conditions of Theorem 5.1. In particular, we have $\{y_{k,0}\xi^j\} \ll Y_k^{-1}$ for $k \geqslant 1$ and $j = 0, 1, 2$. Here, we prove a lower bound for $\{y_{k,0}\xi^3\}$ which implies the measure of approximation of Theorem 1.5 through the following proposition.

PROPOSITION 9.1. *Assume that there exist real numbers δ and c_1 with $0 \leqslant \delta < 1$ and $c_1 > 0$ such that*

$$\{y_{k,0}\xi^3\} \geqslant c_1 Y_k^{-\delta} \tag{9.1}$$

for all $k \geqslant 1$. Then there exists a constant $c_2 > 0$ such that, for any algebraic integer $\alpha \in \mathbb{C}$ of degree at most 3, we have

$$|\xi-\alpha| \geqslant c_2 H(\alpha)^{-\theta}$$

where $\theta = (\gamma^2 + \delta/\gamma)/(1-\delta)$.

Proof. Let α be an algebraic integer of degree at most 3 and let $P(T) = T^3 + pT^2 + qT + r$ be the product of its irreducible polynomial over $\mathbb{Z}$ by the appropriate power of T which makes it of degree 3. Since $H(P) = H(\alpha)$ we find, for all $k \geqslant 1$, that

$$\begin{aligned}\{y_{k,0}\xi^3\} &\leqslant |y_{k,0}P(\xi)| + |p|\{y_{k,0}\xi^2\} + |q|\{y_{k,0}\xi\}\\ &\leqslant c_3(Y_k H(\alpha)|\xi - \alpha| + Y_k^{-1}H(\alpha))\end{aligned}$$

with a constant $c_3 > 0$ depending only on ξ. Choosing k to be the smallest positive integer for which

$$H(\alpha) \leqslant \frac{c_1}{2c_3} Y_k^{1-\delta},$$

and using (9.1) we see that this gives

$$|\xi - \alpha| \geqslant \frac{c_1}{2c_3} Y_k^{-1-\delta} H(\alpha)^{-1}.$$

The conclusion follows since the above choice of k implies that $Y_k \ll H(\alpha)^{\gamma/(1-\delta)}$. □

Numerical experiments done on Fibonacci continued fractions suggest that the values of $\{y_{k,0}\xi^3\}$ are more or less uniformly distributed in the interval $(0, \frac{1}{2})$. If true, this would imply a lower bound of the form $\{y_{k,0}\xi^3\} \gg 1/k \gg (\log\log Y_k)^{-1}$. To be safer, we may hope for a lower bound of the type $\{y_{k,0}\xi^3\} \gg Y_k^{-\delta}$ for any $\delta > 0$. Then, for any $\epsilon > 0$, the above proposition gives a measure of approximation by algebraic integers α of degree at most 3 of the form $|\xi - \alpha| \geqslant cH(\alpha)^{-\gamma^2-\epsilon}$ with a constant c depending on ξ and ϵ. Here we content ourselves with the following much weaker lower bound which, in the same way, implies Theorem 1.5.

PROPOSITION 9.2. *In the above notation, we have* $\{y_{k,0}\xi^3\} \gg Y_k^{-1/\gamma^3}$ *for all sufficiently large values of* k *(to avoid those indices with* $y_{k,0} = 0$*).*

Proof. Fix an integer $k \geqslant 4$ and denote by $y_{k,3}$ the nearest integer to $y_{k,2}\xi$. Define

$$\begin{aligned}\tilde{\mathbf{y}}_k &= (y_{k,1}, y_{k,2}, y_{k,3}),\\ d_{k-2} &= \det(\mathbf{y}_{k-2}, \mathbf{y}_{k-1}, \mathbf{y}_k),\\ \tilde{d}_{k-2} &= \det(\mathbf{y}_{k-2}, \mathbf{y}_{k-1}, \tilde{\mathbf{y}}_k).\end{aligned}$$

According to Theorem 5.1, d_{k-2} belongs to a finite set of non-zero integers. Since ξ is irrational and since $\tilde{d}_{k-2}$ is an integer, we deduce that

$$1 \ll |d_{k-2}\xi - \tilde{d}_{k-2}|.$$

On the other hand, using Proposition 8.1, we find that

$$\begin{aligned}|d_{k-2}\xi - \tilde{d}_{k-2}| &= |\det(\mathbf{y}_{k-2}, \mathbf{y}_{k-1}, \xi\mathbf{y}_k - \tilde{\mathbf{y}}_k)|\\ &\ll \|\xi\mathbf{y}_k - \tilde{\mathbf{y}}_k\| H(Q_{k-2})\\ &\ll (\{y_{k,2}\xi\} + Y_k^{-1})Y_{k-3}.\end{aligned}$$

Combining these two estimates, we obtain $\{y_{k,2}\xi\} \gg Y_{k-3}^{-1} \gg Y_k^{-1/\gamma^3}$ for all sufficiently large k. The conclusion follows since $|y_{k,0}\xi^3 - y_{k,2}\xi| \ll Y_k^{-1}$. □

10. *A partial converse*

We conclude this paper with the following partial converse to Proposition 9.1, where the notation is as in § 9.

PROPOSITION 10.1. *Let ℓ be the least common multiple of all integers of the form $d_k = \det(\mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2})$ with $k \geqslant 1$, and assume that there exist infinitely many indices $k \geqslant 1$ satisfying*

$$\left\{\frac{y_{k,0}\xi^3}{\ell}\right\} \leqslant c_1 Y_k^{-\delta} \tag{10.1}$$

for some real numbers $c_1 > 0$ and $\delta > 0$. Then there exist a constant $c_2 > 0$ and infinitely many algebraic integers $\alpha \in \mathbb{C}$ of degree at most 3 for which

$$|\xi - \alpha| \leqslant c_2 H(\alpha)^{-\theta}$$

where $\theta = (\gamma^2 - \delta)/(1 - \delta)$.

Our proof below follows by making explicit the arguments of Davenport and Schmidt in [**5**, § 2]. To make the connection clearer for the interested reader, denote by $\mathcal{C}_k$ the convex body of $\mathbb{R}^3$ defined by

$$|x_0| \leqslant Y_k^{\gamma}, \quad |x_0\xi - x_1| \leqslant Y_k^{-1} \quad \text{and} \quad |x_0\xi^2 - x_2| \leqslant Y_k^{-1}.$$

Then there exists a constant $c > 0$ which is independent of k such that $\mathbf{y}_{k+1} \in c\,\mathcal{C}_k$, $\mathbf{y}_k \in c\,\mathcal{C}_k$ and $\mathbf{y}_{k-1} \in cY_k^{1/\gamma^2}\mathcal{C}_k$ for $k \geqslant 2$. Since the volume of $\mathcal{C}_k$ is $8Y_k^{-1/\gamma^2}$ and since $\mathbf{y}_{k+1}$, $\mathbf{y}_k$ and $\mathbf{y}_{k-1}$ are linearly independent over $\mathbb{Q}$, this implies that the successive minima of $\mathcal{C}_k$ behave like 1, 1 and Y_k^{1/γ^2} and that they are essentially realized by these three points. Then, the successive minima of the polar convex body $\mathcal{C}_k^*$ defined by

$$|x_0 + x_1\xi + x_2\xi^2| \leqslant Y_k^{-\gamma}, \quad |x_1| \leqslant Y_k \quad \text{and} \quad |x_2| \leqslant Y_k$$

behave like Y_k^{-1/γ^2}, 1 and 1 and, by identifying a point $(x_0, x_1, x_2) \in \mathbb{Z}^3$ with the polynomial $x_0 + x_1T + x_2T^2 \in \mathbb{Z}[T]$, these minima are essentially realized by the polynomials denoted B, A and C in the proof below.

Proof. Fix an index $k \geqslant 2$ for which (10.1) is satisfied. Define $\mathbf{T} = (1, T, T^2)$ and consider the three polynomials

$$A = \det(\mathbf{T}, \mathbf{y}_{k-1}, \mathbf{y}_{k+1}), \quad B = \det(\mathbf{T}, \mathbf{y}_k, \mathbf{y}_{k+1}) \quad \text{and} \quad C = \det(\mathbf{T}, \mathbf{y}_{k-1}, \mathbf{y}_k).$$

We find that

$$d_{k-1} = y_{k,0}A(T) + y_{k-1,0}B(T) + y_{k+1,0}C(T).$$

Thus, if we put $m = \ell/d_{k-1}$, then, for a suitable choice of signs, the polynomial

$$\begin{aligned} P(T) = \; & T^3 - \xi^3 \pm m\left\{\frac{y_{k,0}\xi^3}{\ell}\right\}A(T) \\ & \pm m\left\{\frac{y_{k-1,0}\xi^3}{\ell}\right\}B(T) \pm m\left\{\frac{y_{k+1,0}\xi^3}{\ell}\right\}C(T) \end{aligned}$$

has integral coefficients and is monic of degree 3. We claim that it also satisfies

$$|P'(\xi)| \sim H(P) \quad \text{and} \quad |P(\xi)| \ll H(P)^{-\gamma/(1-\delta)},$$

and that $H(P)$ tends to infinity with k. If we take this for granted, then the root α of P which is closest to ξ is an algebraic integer of degree at most 3 with

$$|\xi - \alpha| \ll \frac{|P(\xi)|}{|P'(\xi)|} \ll H(P)^{-\theta} \ll H(\alpha)^{-\theta}$$

and the proposition is proved.

To establish the claim, we first note that, by (10.1) and Proposition 9.2, we have

$$Y_k^{-\delta} \gg \left\{\frac{y_{k,0}\xi^3}{\ell}\right\} \geqslant \frac{1}{\ell}\{y_{k,0}\xi^3\} \gg Y_k^{-1/\gamma^3}, \tag{10.2}$$

and so $\delta \leqslant \gamma^{-3}$ since k may be arbitrarily large. We also observe that, in the notation of Proposition 8.1, we have $B = Q_k$ and $C = Q_{k-1}$. Thus, with the appropriate choice of sign, we get

$$H\left(P \pm m\left\{\frac{y_{k,0}\xi^3}{\ell}\right\}A\right) \ll 1 + H(B) + H(C) \ll Y_{k-1} \tag{10.3}$$

and

$$|P(\xi)| \ll |A(\xi)| + |B(\xi)| + |C(\xi)| \ll |A(\xi)| + Y_{k+1}^{-1}. \tag{10.4}$$

Moreover, by Corollary 5.3, the point $d_{k-1}\mathbf{y}_{k+2} - d_k\mathbf{y}_{k-1}$ is an integral multiple of $\mathbf{y}_{k+1}$ and thus we find that

$$d_{k-1}Q_{k+1} + d_kA = \det(\mathbf{T}, \mathbf{y}_{k+1}, d_{k-1}\mathbf{y}_{k+2} - d_k\mathbf{y}_{k-1}) = 0.$$

Since d_{k-1} and d_k are non-zero integers of bounded absolute value, this implies, by Proposition 8.1, that

$$H(A) \sim |A'(\xi)| \sim Y_k \quad \text{and} \quad |A(\xi)| \ll Y_{k+3}^{-1}. \tag{10.5}$$

Combining (10.2), (10.3) and (10.5) we deduce that

$$H(P) \sim \left\{\frac{y_{k,0}\xi^3}{\ell}\right\}Y_k \sim |P'(\xi)|.$$

In particular this gives $H(P) \ll Y_k^{1-\delta}$ and so, using (10.4) and (10.5), we get

$$|P(\xi)| \ll Y_k^{-\gamma} \ll H(P)^{-\gamma/(1-\delta)}$$

which completes the proof of the claim. □

COROLLARY 10.2. *Let the hypotheses be as in Theorem* 6.2 *and assume moreover that* $\operatorname{trace}(JAB) = \pm 1$. *Then the following conditions are equivalent:*

(i) *there exists* $\theta > \gamma^2$ *such that the inequality* $|\xi - \alpha| \leqslant H(\alpha)^{-\theta}$ *has infinitely many solutions in algebraic integers* $\alpha \in \mathbb{C}$ *of degree at most* 3 *over* $\mathbb{Q}$;

(ii) *there exists* $\delta > 0$ *such that the inequality* $\{y_{k,0}\xi^3\} \leqslant Y_k^{-\delta}$ *holds for infinitely many indices* k.

Proof. The fact that (i) implies (ii) follows from Proposition 9.1. The reverse implication follows from Proposition 10.1 granted that, for the sequence $(\mathbf{y}_k)_{k\geqslant -1}$ constructed by Lemma 6.1, we have $|\det(\mathbf{y}_k, \mathbf{y}_{k+1}, \mathbf{y}_{k+2})| = 1$ for all $k \geqslant -1$. □

In particular, the corollary applies to any Fibonacci continued fraction $\xi_{a,b}$ for which $|a - b| = 1$.

Acknowledgements. The author thanks Drew Vandeth for pointing out the existence of the reference [**1**] and for recognizing the continued fraction expansion of Corollary 6.3 as a Fibonacci word.

References

1. J.-P. Allouche, J. L. Davison, M. Queffélec and L. Q. Zamboni, 'Transcendence of Sturmian or morphic continued fractions', *J. Number Theory* 91 (2001) 39 – 66.
2. Y. Bugeaud and O. Teulié, 'Approximation d'un nombre réel par des nombres algébriques de degré donné', *Acta Arith.* 93 (2000) 77 – 86.
3. J. W. S. Cassels, *An introduction to Diophantine approximation* (Cambridge University Press, 1957).
4. H. Davenport and W. M. Schmidt, 'Approximation to real numbers by quadratic irrationals', *Acta Arith.* 13 (1967) 169 – 176.
5. H. Davenport and W. M. Schmidt, 'Approximation to real numbers by algebraic integers', *Acta Arith.* 15 (1969) 393 – 416.
6. M. Laurent and D. Roy, 'Criteria of algebraic independence with multiplicities and interpolation determinants', *Trans. Amer. Math. Soc.* 351 (1999) 1845 – 1870.
7. M. Lothaire, *Combinatorics on words*, Encyclopedia of Mathematics and its Applications 17 (Addison-Wesley, Reading, MA, 1983).
8. D. Roy, 'Approximation simultanée d'un nombre et de son carré', *C. R. Acad. Sci. Paris Sér. I* 336 (2003) 1 – 6.
9. W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics 785 (Springer, Berlin, 1980).
10. E. Wirsing, 'Approximation mit algebraischen Zahlen beschränkten Grades', *J. reine angew. Math.* 206 (1961) 67 – 77.

Damien Roy
Département de Mathématiques et de Statistiques
Université d'Ottawa
585 King Edward
Ottawa
Ontario K1N 6N5
Canada

droy@uottawa.ca

http://aix1.uottawa.ca/∼droy/