ACTA ARITHMETICA
133.2 (2008)

On simultaneous rational approximations to a real number, its square, and its cube

by

Damien Roy (Ottawa)

Au Professeur Wolfgang Schmidt, avec mes meilleurs vœux et toute mon estime

1. Introduction. In a remarkable paper [3], H. Davenport and W. M. Schmidt showed that, for any integer $n \geq 2$ and for any real number ξ which is not algebraic over $\mathbb{Q}$ of degree at most $n-1$, there exist infinitely many algebraic integers α of degree at most n satisfying

$$|\xi - \alpha| \leq cH(\alpha)^{-\tau(n)},$$

where $c = c(n,\xi) > 0$ is an appropriate constant depending only on n and ξ, and where $\tau(2) = 2$, $\tau(3) = (3+\sqrt{5})/2$, $\tau(4) = 3$ and $\tau(n) = \lfloor (n+1)/2 \rfloor$ if $n \geq 5$. For $n = 2, 3$, this value of $\tau(n)$ cannot be improved (see [3] for the case $n = 2$ and [7] for the case $n = 3$). For $n \geq 4$, M. Laurent showed in [4] that $\tau(n)$ can be taken to be $\lceil (n+1)/2 \rceil$. However, at present, no optimal value for $\tau(n)$ is known for any single value of $n \geq 4$. Furthermore, we possess no non-trivial upper bound for $\tau(n)$ for $n \geq 4$, besides the estimate $\tau(n) \leq n$ coming from metrical considerations (by an application of the Borel–Cantelli lemma as in the proof of [1, Thm. 3.3]). Although we shall not go into this, let us simply mention that the situation is similar in the case of approximation by algebraic numbers of degree at most n. In this case, it is only for $n \leq 2$ that the optimal exponents are known, the case $n = 2$ being due once again to Davenport and Schmidt [2].

Several years ago, I started working on finding an optimal value for $\tau(4)$ (in the above notation) and, in spite of much effort, I was not successful. My hopes were that this would lead to a new class of extremal numbers,

---

2000 *Mathematics Subject Classification*: Primary 11J13; Secondary 11J04.

Key words and phrases: height, algebraic numbers, approximation to real numbers, exponent of approximation, simultaneous approximation.

Work partially supported by NSERC and CICMA.

similar to that of [5] or [6, §6], and that such a construction could be generalized to larger values of n to provide a non-trivial upper bound for the corresponding values of $\tau(n)$, and maybe settle the question as to whether $\limsup_{n\to\infty} \tau(n)/n$ is equal to 1 or strictly smaller than 1. These problems remain open.

The method initiated by Davenport and Schmidt in [3] for estimating $\tau(n)$ is based on geometry of numbers and requires an upper bound on the uniform exponent of simultaneous approximation to the first $n-1$ consecutive powers of a real number ξ by rational numbers with the same denominator. By [3, §2, Lemma 1], our main result below implies that $\tau(4)$ can be taken to be $\lambda_3^{-1} + 1 \cong 3.3556$, where

$$\lambda_3 = \frac{1}{2}\Big(2 + \sqrt{5} - \sqrt{7 + 2\sqrt{5}}\Big) \cong 0.4245.$$

THEOREM. *Let $\xi \in \mathbb{R}$ with $[\mathbb{Q}(\xi) \colon \mathbb{Q}] > 3$, and let c and λ be positive real numbers. Suppose that for any sufficiently large value of X, the inequalities*

$$|x_0| \le X, \quad |x_0\xi - x_1| \le cX^{-\lambda}, \quad |x_0\xi^2 - x_2| \le cX^{-\lambda}, \quad |x_0\xi^3 - x_3| \le cX^{-\lambda}$$

admit a non-zero solution $\mathbf{x} = (x_0, x_1, x_2, x_3) \in \mathbb{Z}^4$. Then $\lambda \le \lambda_3$. Moreover, if $\lambda = \lambda_3$, then c is bounded below by a positive constant depending only on ξ.

The rest of the paper is devoted to the proof of this result, which, through its weaker hypothesis on ξ, complements [3, Theorem 4a]. The tools that we use for the proof are the same as those of [3] together with results on heights of subspaces of $\mathbb{R}^n$ defined over $\mathbb{Q}$ that were developed around the same period of time by W. M. Schmidt in [8]. Using other tools, similar to the bracket $[\mathbf{x}, \mathbf{y}, \mathbf{z}]$ in [6, §2], I discovered recently that the exponent λ_3 in the above theorem is not optimal. Since the argument is quite involved and does not seem to lead to a significant improvement in λ_3, I decided not to include this here.

2. First considerations. Throughout this paper, we fix a real number ξ with $[\mathbb{Q}(\xi) \colon \mathbb{Q}] > 3$ and positive constants λ, c satisfying the hypotheses of the Theorem. In all statements below, the implied constants in the symbols $\gg$, $\ll$ and $\asymp$ (the conjunction of $\gg$ and $\ll$) depend only on ξ and λ (not on c). In particular, we may assume that $c \ll 1$. Our goal is to show that $\lambda \le \lambda_3$ and that $c \gg 1$ in case of equality. By [3, Theorem 4a], we already have $\lambda \le 1/2$.

For each integer $n \ge 1$ and each point $\mathbf{x} = (x_0, x_1, \dots, x_n) \in \mathbb{R}^{n+1}$, we define points $\mathbf{x}^-$ and $\mathbf{x}^+$ of $\mathbb{R}^n$ by

$$\mathbf{x}^- = (x_0, \dots, x_{n-1}) \quad \text{and} \quad \mathbf{x}^+ = (x_1, \dots, x_n).$$

We also put

$$\|\mathbf{x}\| = \max_{0\le i\le n} |x_i| \quad \text{and} \quad L(\mathbf{x}) = \max_{1\le i\le n} |x_0\xi^i - x_i|.$$

Finally, we say that a point $\mathbf{x} \in \mathbb{Z}^{n+1}$ is *primitive* if it is non-zero and if the gcd of its coordinates is 1. Then the hypothesis implies that, for any sufficiently large X, there exists a primitive point $\mathbf{x} \in \mathbb{Z}^4$ with

$$\|\mathbf{x}\| \le X \quad \text{and} \quad L(\mathbf{x}) \le cc_1X^{-\lambda}, \tag{1}$$

where $c_1 = 2\max\{1, |\xi|\}^{3\lambda}$. The following lemmas extend results of Davenport and Schmidt in [3, §4].

LEMMA 2.1. *Let $C \in \mathbb{Z}^2$ and $\mathbf{x} \in \mathbb{Z}^{n+1}$ with $n \in \{1, 2, 3\}$. Then the point $\mathbf{y} = C^+\mathbf{x}^- - C^-\mathbf{x}^+$ satisfies*

$$\|\mathbf{y}\| \le \|\mathbf{x}\|L(C) + c_2\|C\|L(\mathbf{x}) \quad \textit{and} \quad L(\mathbf{y}) \le c_2\|C\|L(\mathbf{x}) \tag{2}$$

for some constant $c_2 = c_2(\xi)$. Moreover, if $\mathbf{y} = 0$ and if C and $\mathbf{x}$ are non-zero and primitive, then

$$\|\mathbf{x}\| = \|C\|^n \quad \textit{and} \quad L(\mathbf{x}) \asymp \|C\|^{n-1}L(C).$$

Proof. Write $C = (a, b)$. Then the estimates in (2) follow respectively from the formulas $\mathbf{y} = (b - a\xi)\mathbf{x}^- + a(\xi\mathbf{x}^- - \mathbf{x}^+)$ and $\mathbf{y} = b\mathbf{x}^- - a\mathbf{x}^+$, upon choosing c_2 so that $\|\xi\mathbf{x}^- - \mathbf{x}^+\| \le c_2L(\mathbf{x})$ and $L(\mathbf{x}^-) + L(\mathbf{x}^+) \le c_2L(\mathbf{x})$. If $\mathbf{y} = 0$ and $C \ne 0$, then $\mathbf{x}$ is a rational multiple of the geometric progression $(a^n, a^{n-1}b, \ldots, b^n)$. If furthermore C and $\mathbf{x}$ are primitive, this progression is a primitive point of $\mathbb{Z}^{n+1}$ and so it coincides with $\pm\mathbf{x}$. This gives $\|\mathbf{x}\| = \|C\|^n$ and $L(\mathbf{x}) \asymp \|\mathbf{x}^+ - \xi\mathbf{x}^-\| = \|C\|^{n-1}L(C)$. ∎

LEMMA 2.2. *Suppose that $\lambda > 1/3$. Then for any non-zero point $C \in \mathbb{Z}^2$ we have $L(C) \gg \|C\|^{-1/\lambda}$.*

Proof. Since $\xi \notin \mathbb{Q}$, we have $L(C) \ne 0$ for any non-zero point $C \in \mathbb{Z}^2$. So, it suffices to prove that $L(C) \gg \|C\|^{-1/\lambda}$ for primitive points $C \in \mathbb{Z}^2$ of sufficiently large norm. Let C be a primitive point of $\mathbb{Z}^2$, and let $\mathbf{x} \in \mathbb{Z}^4$ be a primitive solution of (1) for the choice of $X = (2cc_1c_2\|C\|)^{1/\lambda}$, where c_2 is the constant introduced in Lemma 2.1. Since $\lambda > 1/3$, we have $X < \|C\|^3$ if $\|C\| \gg 1$, and then the second part of Lemma 2.1 shows that $\mathbf{y} = C^+\mathbf{x}^- - C^-\mathbf{x}^+$ is a non-zero point of $\mathbb{Z}^3$. Applying the first part of the same lemma, we deduce that

$$1 \le \|\mathbf{y}\| \le XL(C) + cc_1c_2\|C\|X^{-\lambda} \le XL(C) + 1/2,$$

and so $L(C) \ge (2X)^{-1} \gg \|C\|^{-1/\lambda}$. ∎

LEMMA 2.3. *Suppose that $\lambda > 1/3$. Then there exist at most finitely many points $\mathbf{x} \in \mathbb{Z}^4$ with $L(\mathbf{x}) \le cc_1\|\mathbf{x}\|^{-\lambda}$ such that $\mathbf{x}^-$ and $\mathbf{x}^+$ are linearly dependent over $\mathbb{Q}$.*

Proof. Suppose on the contrary that the conclusion is false. Then there exist infinitely many primitive points $\mathbf{x}$ of $\mathbb{Z}^4$ with $L(\mathbf{x}) \le cc_1\|\mathbf{x}\|^{-\lambda}$ for which $\mathbf{x}^-$ and $\mathbf{x}^+$ are linearly dependent. For each of them, there exists a primitive point $C \in \mathbb{Z}^2$ such that $C^+\mathbf{x}^- - C^-\mathbf{x}^+ = 0$. By Lemma 2.1, we have $\|\mathbf{x}\| = \|C\|^3$ and $L(\mathbf{x}) \asymp \|C\|^2 L(C)$. Thus $\|C\|$ tends to infinity with $\|\mathbf{x}\|$, and the condition $L(\mathbf{x}) \le cc_1\|\mathbf{x}\|^{-\lambda}$ translates into $L(C) \ll \|C\|^{-2-3\lambda}$. Since $-2-3\lambda < -3 < -1/\lambda$, this contradicts Lemma 2.2. ∎

LEMMA 2.4. *Let $n \in \{1,2,3\}$ and let U be a proper subspace of $\mathbb{R}^{n+1}$ defined over $\mathbb{Q}$. Then the function $L(\mathbf{x})$ is bounded from below by a positive constant on the set of all non-zero points $\mathbf{x}$ of $U \cap \mathbb{Z}^{n+1}$.*

Proof. As in the proof of [3, §3, Lemma 5], suppose on the contrary that there exists a sequence of non-zero integral points $(\mathbf{x}_i)_{i\ge1}$ in U such that $\lim_{i\to\infty} L(\mathbf{x}_i) = 0$. Then, for any sufficiently large index i, the first coordinate $x_{i,0}$ of $\mathbf{x}$ is non-zero and the product $x_{i,0}^{-1}\mathbf{x}_i$ converges to $(1,\xi,\dots,\xi^n)$ as i tends to infinity. Thus, the point $(1,\xi,\dots,\xi^n)$ belongs to U. This is impossible since U is a proper subspace of $\mathbb{R}^{n+1}$ defined over $\mathbb{Q}$ while the coordinates of the point $(1,\xi,\dots,\xi^n)$ are linearly independent over $\mathbb{Q}$. ∎

Finally, we note that there exists a sequence of non-zero points $(\mathbf{x}_i)_{i\ge1}$ in $\mathbb{Z}^4$ with the following properties:

(a) the positive integers $X_i := \|\mathbf{x}_i\|$ form a strictly increasing sequence,
(b) the positive real numbers $L_i := L(\mathbf{x}_i)$ form a strictly decreasing sequence,
(c) if some non-zero point $\mathbf{x} \in \mathbb{Z}^4$ satisfies $L(\mathbf{x}) < L_i$ for some $i \ge 1$, then $\|\mathbf{x}\| \ge X_{i+1}$.

We fix such a choice of sequence $(\mathbf{x}_i)_{i\ge1}$ and refer to it as the sequence of *minimal points* for ξ although it is not unique and differs from the notion introduced by Davenport and Schmidt in [3, §4]. We note that, for each $i \ge 1$, $\mathbf{x}_i$ is a primitive point of $\mathbb{Z}^4$ and, since (1) admits a non-zero solution $\mathbf{x} \in \mathbb{Z}^4$ for each X with $X_i \le X < X_{i+1}$ when i is sufficiently large, we deduce from condition (c) that

$$L_i \le cc_1 X_{i+1}^{-\lambda}$$

for each large enough index i. We will use this property repeatedly in what follows, either in this form or in the weaker form $L_i \ll cX_{i+1}^{-\lambda} \ll X_{i+1}^{-\lambda}$.

3. A family of planes in $\mathbb{R}^4$. For each integer $n \ge 1$ and each subspace S of $\mathbb{R}^n$ defined over $\mathbb{Q}$ of dimension $p > 0$, we define the *height* $H(S)$ of S by $H(S) = \|\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_p\|$, where $(\mathbf{y}_1,\dots,\mathbf{y}_p)$ is a basis of the group $S \cap \mathbb{Z}^n$ of integral points of S (upon identifying $\bigwedge^p \mathbb{R}^n$ with $\mathbb{R}^{\binom{n}{p}}$ through an ordering of the Grassmann coordinates, as in [9, Chap. 1, §5]). We also define $H(0) = 1$. It then follows from [9, Chap. 1, Lemma 8A] that, for any pair of subspaces S and T of $\mathbb{R}^n$ defined over $\mathbb{Q}$, we have

$$H(S \cap T)H(S+T) \le c(n)H(S)H(T) \tag{3}$$

with a constant $c(n) > 0$ depending only on n. We also recall the duality formula $H(S) = H(S^{\perp})$ where $S^{\perp}$ stands for the orthogonal complement of S in $\mathbb{R}^n$ (see [9, Chap. 1, §8]).

For each $i \ge 2$, we denote by W_i the subspace of $\mathbb{R}^4$ of dimension 2 generated by $\mathbf{x}_{i-1}$ and $\mathbf{x}_i$. We also introduce a new parameter

$$\theta = \frac{1-\lambda}{\lambda},$$

and note that $\theta \ge 1$ since $\lambda \le 1/2$.

Lemma 3.1. *For each $i \ge 2$, the points $\mathbf{x}_{i-1}$ and $\mathbf{x}_i$ form a basis of $W_i \cap \mathbb{Z}^4$, and we have $H(W_i) \asymp X_i L_{i-1} \ll X_i^{1-\lambda}$.*

This follows by a simple adaptation of the proofs of [2, Lemma 2] and [6, Lemma 4.1], the difference being that here X_i stands for the norm of $\mathbf{x}_i$ instead of the absolute value of its first coordinate. We now look at the sums $W_i + W_{i+1}$.

Lemma 3.2. *There exist infinitely many indices $i \ge 2$ such that $W_i \neq W_{i+1}$. For each of them, we have*

$$H(W_i + W_{i+1}) \ll X_i^{-1} H(W_i)H(W_{i+1}) \ll H(W_i)^{-1/\theta} H(W_{i+1}). \tag{4}$$

Proof. If there were only finitely many $i \ge 2$ for which $W_i \neq W_{i+1}$, then all points $\mathbf{x}_i$ with i sufficiently large would lie in a fixed subspace W of $\mathbb{R}^4$ defined over $\mathbb{Q}$ of dimension 2, contrary to Lemma 2.4. This proves the first assertion of the present lemma.

Applying (3) with $S = W_i$ and $T = W_{i+1}$, we find

$$H(W_i \cap W_{i+1})H(W_i + W_{i+1}) \ll H(W_i)H(W_{i+1}).$$

For each index $i \ge 2$ such that $W_i \neq W_{i+1}$, we have $W_i \cap W_{i+1} = \langle \mathbf{x}_i \rangle_{\mathbb{R}}$ and so $H(W_i \cap W_{i+1}) = X_i$. This leads to the first estimate in (4). For the second one, we simply use the lower bound $X_i \gg H(W_i)^{1/(1-\lambda)}$ coming from Lemma 3.1. ∎

Notation. We denote by I the set of indices $i \ge 2$ for which $W_i \neq W_{i+1}$, ordered by increasing magnitude.

Thus, for each $i \in I$, the sum $W_i + W_{i+1} = \langle \mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1} \rangle_{\mathbb{R}}$ is a three-dimensional subspace of $\mathbb{R}^4$ defined over $\mathbb{Q}$. By Lemma 2.4 such a subspace of $\mathbb{R}^4$ contains at most finitely many minimal points. This leads to the first assertion of the next lemma.

LEMMA 3.3. *There exist infinitely many pairs of consecutive elements i, j of I with $i < j$ and $W_i + W_{i+1} \neq W_j + W_{j+1}$. For any such pair of integers i and j, we have*

$$X_i X_j \ll H(W_i)H(W_j)H(W_{j+1}), \tag{5}$$

$$H(W_i)H(W_j) \ll H(W_{j+1})^\theta \quad \text{and} \quad X_i X_j \ll X_{j+1}^\theta. \tag{6}$$

Proof. For consecutive elements $i < j$ of I, we have $W_i \neq W_{i+1} = W_j \neq W_{j+1}$. If $W_i + W_{i+1}$ and $W_j + W_{j+1}$ are distinct subspaces of $\mathbb{R}^4$, their sum is the whole of $\mathbb{R}^4$ and their intersection is $W_{i+1} = W_j$. Since $H(\mathbb{R}^4) = 1$, we deduce from (3) that

$$H(W_{i+1}) \ll H(W_i + W_{i+1})H(W_j + W_{j+1}).$$

Combining this estimate with the upper bounds

$$H(W_i + W_{i+1}) \ll X_i^{-1} H(W_i)H(W_{i+1}),$$
$$H(W_j + W_{j+1}) \ll X_j^{-1} H(W_j)H(W_{j+1})$$

provided by Lemma 3.2, we obtain (5). Then combining (5) with the standard upper bounds $H(W_i) \ll X_i^{1-\lambda}$ and $H(W_j) \ll X_j^{1-\lambda}$ coming from Lemma 3.1, we find

$$X_i^\lambda X_j^\lambda \ll H(W_{j+1}),$$

so $H(W_i)H(W_j) \ll (X_i X_j)^{1-\lambda} \ll H(W_{j+1})^\theta \ll X_{j+1}^{\theta(1-\lambda)}$, which proves (6). ∎

4. A family of points in $\mathbb{Z}^2$. For each pair of points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{Z}^4$, we define

$$C(\mathbf{x}, \mathbf{y}) = (\det(\mathbf{x}^-, \mathbf{x}^+, \mathbf{y}^-), \det(\mathbf{x}^-, \mathbf{x}^+, \mathbf{y}^+)) \in \mathbb{Z}^2.$$

To alleviate the notation, we also write

$$C_{i,j} = C(\mathbf{x}_i, \mathbf{x}_j)$$

for each pair of integers $i, j \geq 1$. These points $C_{i,j}$ play a crucial role in the proof of the inequality $\lambda \leq 1/2$ by Davenport and Schmidt in [3, §4]. They also play an important role in the present work. We first prove general estimates.

LEMMA 4.1. *For any pair of integers $i, j \geq 1$, we have*

$$\|C_{i,j}\| \ll X_j L_i^2 + X_i L_i L_j \quad \text{and} \quad L(C_{i,j}) \ll X_i L_i L_j.$$

Proof. The estimate for $\|C_{i,j}\|$ is standard (see for example the proof of [3, §4, Lemma 7]). For the other quantity, we find

$$\begin{aligned} L(C_{i,j}) &= |\det(\mathbf{x}_i^-, \mathbf{x}_i^+, \mathbf{x}_j^+ - \xi\mathbf{x}_j^-)| \\ &= |\det(\mathbf{x}_i^-, \mathbf{x}_i^+ - \xi\mathbf{x}_i^-, \mathbf{x}_j^+ - \xi\mathbf{x}_j^-)| \ll X_i L_i L_j. \ \blacksquare \end{aligned}$$

The next lemma provides a sharper upper bound for $L(C_{i,i+1})$ when $i \in I$.

LEMMA 4.2. *Let $i < j$ be consecutive elements of the set I. Then $C_{i,j} = bC_{i,i+1}$ for some non-zero integer b with $|b| \asymp X_j/X_{i+1}$, and we have*

$$L(C_{i,i+1}) \ll X_i X_j^{-\lambda} X_{j+1}^{-\lambda}.$$

Proof. Since i and j are consecutive in I, we have $W_{i+1} = W_j$. Moreover, since $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ form a basis of the group of integral points of W_{i+1}, there exist integers a and b with $b \neq 0$ such that $\mathbf{x}_j = a\mathbf{x}_i + b\mathbf{x}_{i+1}$. If $X_j > 3|b|X_{i+1}$, we deduce that

$$|a|X_i = \|\mathbf{x}_j - b\mathbf{x}_{i+1}\| \geq X_j - |b|X_{i+1} > 2|b|X_{i+1},$$

and so $|a| > 2|b|$. Then, we find $L_j \geq |a|L_i - |b|L_{i+1} > |b|L_{i+1} \geq L_{i+1}$, which is impossible. This contradiction shows that $|b| \geq X_j/(3X_{i+1})$. Since the point $C(\mathbf{x}, \mathbf{y})$ is a linear function of $\mathbf{y}$ and since $C(\mathbf{x}, \mathbf{x}) = 0$ for any $\mathbf{x} \in \mathbb{R}^4$, we also have

$$C_{i,j} = C(\mathbf{x}_i, a\mathbf{x}_i + b\mathbf{x}_{i+1}) = bC_{i,i+1}$$

and so, by Lemma 4.1, we obtain (since $\lambda \leq 1/2 \leq 1$)

$$L(C_{i,i+1}) = |b|^{-1}L(C_{i,j}) \leq |b|^{-\lambda}L(C_{i,j}) \ll \frac{X_{i+1}^{\lambda}}{X_j^{\lambda}} X_i L_i L_j \ll X_i X_j^{-\lambda} X_{j+1}^{-\lambda}. \blacksquare$$

REMARK. Although we will not use this here, it is interesting to note that the identity

$$\det(\mathbf{w}, \mathbf{x}, \mathbf{y})\mathbf{z} - \det(\mathbf{w}, \mathbf{x}, \mathbf{z})\mathbf{y} + \det(\mathbf{w}, \mathbf{y}, \mathbf{z})\mathbf{x} - \det(\mathbf{x}, \mathbf{y}, \mathbf{z})\mathbf{w} = 0,$$

which holds for any quadruple of points $(\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z})$ in $\mathbb{R}^3$, specializes to

$$C_{i,j}^{+}\mathbf{x}_j^{-} - C_{i,j}^{-}\mathbf{x}_j^{+} = C_{j,i}^{-}\mathbf{x}_i^{+} - C_{j,i}^{+}\mathbf{x}_i^{-}$$

when we apply it to the quadruple $(\mathbf{x}_i^{-}, \mathbf{x}_i^{+}, \mathbf{x}_j^{-}, \mathbf{x}_j^{+})$ for a choice of integers $i, j \geq 1$.

5. A family of planes in $\mathbb{R}^3$. From now on, we assume that $\lambda > 1/3$. Then, by Lemma 2.3, there exists an index i_0 such that $\mathbf{x}_i^{-}$ and $\mathbf{x}_i^{+}$ are linearly independent for each $i \geq i_0$. For those values of i, we denote by V_i the two-dimensional subspace of $\mathbb{R}^3$ spanned by these points:

$$V_i = \langle \mathbf{x}_i^{-}, \mathbf{x}_i^{+} \rangle_{\mathbb{R}}.$$

Since $\max\{L(\mathbf{x}_j^{-}), L(\mathbf{x}_j^{+})\} \ll L_j$ tends to 0 as $j \to \infty$, it follows from Lemma 2.4 that each V_i contains at most finitely many points of the form $\mathbf{x}_j^{-}$ or $\mathbf{x}_j^{+}$, and so there are infinitely many indices $i \geq i_0$ such that $V_i \neq V_{i+1}$. We also note that, for $i, j \geq i_0$, we have

$$V_i = V_j \;\Leftrightarrow\; C_{i,j} = 0 \;\Leftrightarrow\; C_{j,i} = 0$$

by definition of the points $C_{i,j}$ (see §4). In [3, §4], Davenport and Schmidt argue that, for each $i \geq i_0$ such that $V_i \neq V_{i+1}$, we have $1 \leq \|C_{i,i+1}\| \ll X_{i+1}L_i^2 \ll X_{i+1}^{1-2\lambda}$ (see Lemma 4.1). Since i can be taken to be arbitrarily large, this gives $1 - 2\lambda \geq 0$ and so $\lambda \leq 1/2$.

LEMMA 5.1. *There exist infinitely many integers $i > i_0$ for which $V_{i-1} \neq V_i$. For each of them, we have*

$$H(W_{i+1}) \ll X_{i+1}^{1-\lambda} \ll H(W_i)^\theta \ll X_i^{\theta(1-\lambda)}. \tag{7}$$

In particular, this leads to the symmetric estimates $X_{i+1} \ll X_i^\theta$ and $H(W_{i+1}) \ll H(W_i)^\theta$.

Proof. The first assertion being already settled, fix an index $i > i_0$ such that $V_{i-1} \neq V_i$. Then the integral point $C_{i,i-1}$ is non-zero and so its norm is bounded below by 1. The absolute values of its coordinates are:

$$|\det(\mathbf{x}_i^-, \mathbf{x}_i^+, \mathbf{x}_{i-1}^-)| = |\det(\mathbf{x}_{i-1}^-, \mathbf{x}_i^-, \mathbf{x}_i^+ - \xi\mathbf{x}_i^-)| \ll \|\mathbf{x}_{i-1}^- \wedge \mathbf{x}_i^-\| L_i,$$
$$|\det(\mathbf{x}_i^-, \mathbf{x}_i^+, \mathbf{x}_{i-1}^+)| = |\det(\mathbf{x}_{i-1}^+, \mathbf{x}_i^+, \mathbf{x}_i^- - \xi^{-1}\mathbf{x}_i^+)| \ll \|\mathbf{x}_{i-1}^+ \wedge \mathbf{x}_i^+\| L_i.$$

Since $\|\mathbf{x}_{i-1}^- \wedge \mathbf{x}_i^-\|$ and $\|\mathbf{x}_{i-1}^+ \wedge \mathbf{x}_i^+\|$ are bounded above by $\|\mathbf{x}_{i-1} \wedge \mathbf{x}_i\| = H(W_i)$, this means that $\|C_{i,i-1}\| \ll H(W_i)L_i$. Thus we obtain

$$1 \leq \|C_{i,i-1}\| \ll H(W_i)L_i \ll H(W_i)X_{i+1}^{-\lambda},$$

and so $X_{i+1} \ll H(W_i)^{1/\lambda}$. The conclusion follows by combining this result with the estimates $H(W_i) \ll X_i^{1-\lambda}$ and $H(W_{i+1}) \ll X_{i+1}^{1-\lambda}$ coming from Lemma 3.1. ∎

PROPOSITION 5.2. *Suppose that there exist infinitely many indices $i \geq i_0$ such that $V_i = V_{i+1}$. Then $\lambda \leq \sqrt{2} - 1 \cong 0.4142$. Moreover, if $\lambda = \sqrt{2} - 1$, then we also have $c \gg 1$.*

Proof. Since there are infinitely many indices $i > i_0$ for which $V_{i-1} \neq V_i$, the hypothesis of the proposition forces the existence of arbitrarily large indices i with

$$V_{i-1} \neq V_i = V_{i+1}.$$

Fix such an i. Let $px_0 + qx_1 + rx_2 = 0$ be an equation of V_i with relatively prime coefficients $p, q, r \in \mathbb{Z}$, so that by duality $H(V_i) = \|(p, q, r)\|$. For any point $\mathbf{x} = (x_0, x_1, x_2, x_3)$ of W_{i+1}, we have

$$\mathbf{x}^- = (x_0, x_1, x_2) \in \langle \mathbf{x}_i^-, \mathbf{x}_{i+1}^- \rangle_{\mathbb{R}} \quad \text{and} \quad \mathbf{x}^+ = (x_1, x_2, x_3) \in \langle \mathbf{x}_i^+, \mathbf{x}_{i+1}^+ \rangle_{\mathbb{R}},$$

therefore $\mathbf{x}^-$ and $\mathbf{x}^+$ both belong to $V_i + V_{i+1} = V_i$, and so the point $\mathbf{x}$ satisfies

$$px_0 + qx_1 + rx_2 = 0 \quad \text{and} \quad px_1 + qx_2 + rx_3 = 0.$$

This means that the orthogonal complement of W_i in $\mathbb{R}^4$ is the space $\langle (p, q, r, 0), (0, p, q, r) \rangle_{\mathbb{R}}$ and so, applying the duality property of the height again, we find

$$(8)\qquad H(W_{i+1}) = H(\langle (p,q,r,0),(0,p,q,r)\rangle_{\mathbb{R}}) \asymp \|(p,q,r)\|^2 = H(V_i)^2$$

(the relation $H(V_i) \ll H(W_{i+1})^{1/2}$ also follows from [3, Thm. 3] since the equality $V_i = V_{i+1}$ means that (p,q,r) provides a three-term recurrence relation satisfied by both $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$). We now argue as M. Laurent in the proof of [4, Lemma 5]. Define

$$P(T) = p + qT + rT^2 \in \mathbb{Z}[T].$$

For any point $\mathbf{y} = (y_0, y_1, y_2) \in \mathbb{Z}^3$, we have

$$(9)\qquad |(py_0 + qy_1 + ry_2) - y_0 P(\xi)| \le 2H(V_i)L(\mathbf{y}).$$

Applying this estimate to the point $\mathbf{y} = \mathbf{x}_{i+1}^- \in V_i$, we get

$$(10)\qquad X_{i+1}|P(\xi)| \ll H(V_i)L_{i+1}.$$

Since $V_{i-1} \neq V_i$, at least one of the points $\mathbf{x}_{i-1}^-$ or $\mathbf{x}_{i-1}^+$ does not belong to V_i. If $\mathbf{y} = (y_0, y_1, y_2)$ is such a point, then $py_0 + qy_1 + ry_2$ is a non-zero integer, and using successively (9), (10) and (8) we obtain

$$\begin{aligned} 1 \le |py_0 + qy_1 + ry_2| &\ll X_{i-1}|P(\xi)| + H(V_i)L_{i-1} \ll H(V_i)L_{i-1} \\ &\ll cH(W_{i+1})^{1/2} X_i^{-\lambda}. \end{aligned}$$

Moreover, Lemma 5.1 gives $H(W_{i+1}) \ll X_i^{\theta(1-\lambda)}$ and so the last estimate leads to

$$1 \ll cX_i^{(1-\lambda)^2/(2\lambda)-\lambda} = cX_i^{(2-(1+\lambda)^2)/(2\lambda)}.$$

As i can be taken to be arbitrarily large, this implies that $2-(1+\lambda)^2 \ge 0$, and so $\lambda \le \sqrt{2}-1$. Moreover, we obtain $c \gg 1$ if $\lambda = \sqrt{2}-1$. ∎

Corollary 5.3. *Suppose that $\lambda > \sqrt{2}-1$. Then we have $V_{i-1} \neq V_i$ for any sufficiently large integer i, and the estimates* (7) *of Lemma* 5.1 *apply to all integers $i \ge 1$. Moreover, for any pair of consecutive integers $i < j$ of I with $W_i + W_{i+1} \neq W_j + W_{j+1}$, we also have*

$$(11)\qquad H(W_i) \ll X_i^{1-\lambda} \ll H(W_j)^{\theta^2-1} \ll X_j^{(\theta^2-1)(1-\lambda)},$$

$$(12)\qquad H(W_j) \ll X_j^{1-\lambda} \ll H(W_{j+1})^{\theta(1-\lambda)} \ll X_{j+1}^{\theta(1-\lambda)^2}.$$

Proof. The first assertion follows directly from Lemma 5.1 and the above proposition. To prove the second one, we fix consecutive integers $i < j$ in I with $W_i + W_{i+1} \neq W_j + W_{j+1}$, and go back to the general estimate (5) from Lemma 3.3:

$$(13)\qquad X_i X_j \ll H(W_i)H(W_j)H(W_{j+1}).$$

On the right hand side of this inequality, we apply the standard estimate $H(W_i) \ll X_i^{1-\lambda}$ from Lemma 3.1 as an upper bound for $H(W_i)$, and the estimate $H(W_{j+1}) \ll H(W_j)^{\theta}$ coming from (7) as an upper bound for $H(W_{j+1})$.

On the left hand side, we use instead the estimate $H(W_j) \ll X_j^{1-\lambda}$ from Lemma 3.1 as a lower bound for X_j. This gives

$$X_i^{\lambda} \ll H(W_j)^{\theta+1-1/(1-\lambda)} = H(W_j)^{\theta-1/\theta},$$

and (11) follows. To prove (12), we note instead that, i and j being consecutive elements of I, we have $W_j = W_{i+1}$ and so (13) combined with Lemma 3.1 gives

$$X_iX_j \ll H(W_i)H(W_{i+1})H(W_{j+1}) \ll (X_iX_{i+1})^{1-\lambda}H(W_{j+1}).$$

Moving all powers of X_i to the left and using the estimate $X_{i+1} \ll X_i^{\theta}$ from (7) as a lower bound for X_i, we obtain

$$X_{i+1}^{\lambda/\theta}X_j \ll X_{i+1}^{1-\lambda}H(W_{j+1}).$$

Moving all powers of X_{i+1} to the right and observing that the exponent $1-\lambda-\lambda/\theta = 1-1/\theta$ is ≥ 0 (since $\theta \geq 1$), we finally obtain

$$X_j \ll X_{i+1}^{1-1/\theta}H(W_{j+1}) \leq X_j^{1-1/\theta}H(W_{j+1}),$$

which implies (12). ∎

6. The set J. We assume from now on that $\lambda > \sqrt{2}-1$. Then, for each sufficiently large index i, the subspace $V_i = \langle \mathbf{x}_i^-, \mathbf{x}_i^+\rangle_{\mathbb{R}}$ of $\mathbb{R}^3$ has dimension 2 and, by Corollary 5.3, we have $V_i \neq V_{i+1}$. Consequently, $C_{i,i+1}$ is a non-zero point of $\mathbb{Z}^2$ for each $i \gg 1$.

Notation. Let J be the set of all elements i of I whose successor j in I satisfies $W_j + W_{j+1} \neq W_i + W_{i+1}$.

By Lemma 3.3, the set J is infinite. The next result studies a possible configuration of points.

Lemma 6.1. *Suppose that $\lambda > \sqrt{2}-1$, and that $h < i < j$ are three consecutive elements of I with $h \in J$ and $i \in J$. Then we have*

$$L(C_{i,i+1}) \ll X_{j+1}^{\alpha} \quad \textit{where} \quad \alpha = \frac{-\lambda^4+\lambda^3+\lambda^2-3\lambda+1}{\lambda(\lambda^2-\lambda+1)}.$$

Proof. By Lemma 4.2,

$$L(C_{i,i+1}) \ll X_iX_j^{-\lambda}X_{j+1}^{-\lambda}. \tag{14}$$

Since $i \in J$, we have $W_i + W_{i+1} \neq W_j + W_{j+1}$, and the second part of (6) in Lemma 3.3 gives

$$X_i \ll X_j^{-1}X_{j+1}^{\theta}.$$

Since $h \in J$, we also have $W_h + W_{h+1} \neq W_i + W_{i+1}$, and the estimates (12) of Corollary 5.3 applied to the pair (h,i) instead of (i,j) lead to

$$X_i \ll X_{i+1}^{(1-\lambda)\theta} \leq X_j^{(1-\lambda)\theta}.$$

Put $\beta = (1-\lambda)/(\lambda^2-\lambda+1)$. Since $\lambda \le 1/2$, we have $\beta \ge 1-\lambda \ge 1/2$. We consider two cases.

(a) If $X_j \ge X_{j+1}^{\beta}$, we substitute into (14) the first of the above two upper bounds for X_i. This gives

$$L(C_{i,i+1}) \ll X_j^{-1-\lambda} X_{j+1}^{\theta-\lambda} \le X_{j+1}^{-(1+\lambda)\beta+\theta-\lambda} = X_{j+1}^{\alpha}.$$

(b) If on the contrary, we have $X_j < X_{j+1}^{\beta}$, we substitute instead into (14) the second upper bound for X_i. Again we find

$$L(C_{i,i+1}) \ll X_j^{(1-\lambda)\theta-\lambda} X_{j+1}^{-\lambda} \le X_{j+1}^{((1-\lambda)\theta-\lambda)\beta-\lambda} = X_{j+1}^{\alpha},$$

upon noting that the exponent $(1-\lambda)\theta-\lambda = (1-2\lambda)/\lambda$ is ≥ 0. ■

PROPOSITION 6.2. *Suppose that* $\lambda > \lambda_2$ *where* $\lambda_2 \cong 0.4241$ *denotes the positive root of the polynomial* $P_2(T) = 3T^4 - 4T^3 + 2T^2 + 2T - 1$, *and let* α *be as in Lemma* 6.1. *Then we have* $1 - 2\lambda + \alpha < 0$ *and*, *for any triple of consecutive elements* $h < i < j$ *of* I *contained in* J, *with* i *large enough*, *the points* $C_{i,i+1}$ *and* $C_{j,j+1}$ *are linearly dependent over* $\mathbb{Q}$.

The fact that $P_2(T)$ admits exactly one positive root λ_2 follows by observing that its second derivative $P_2''(T) = (6T-2)^2$ is non-negative on $\mathbb{R}$ and that $P_2(0)$ is negative. Consequently, if $\lambda > \lambda_2$, we have $P_2(\lambda) > 0$.

Proof. For any triple of consecutive elements $h < i < j$ of I contained in J, Lemma 6.1 gives $L(C_{i,i+1}) \ll X_{j+1}^{\alpha}$ and $L(C_{j,j+1}) \ll X_{k+1}^{\alpha}$, where k denotes the successor of j in I. As the general estimates of Lemma 4.1 provide $\|C_{l,l+1}\| \ll X_{l+1}^{1-2\lambda}$ for each $l \ge 1$, we deduce that

$$\begin{aligned}|\det(C_{i,i+1}, C_{j,j+1})| &\ll \|C_{i,i+1}\| L(C_{j,j+1}) + \|C_{j,j+1}\| L(C_{i,i+1}) \\ &\ll X_{i+1}^{1-2\lambda} X_{k+1}^{\alpha} + X_{j+1}^{1-2\lambda+\alpha} \ll X_{k+1}^{1-2\lambda+\alpha} + X_{j+1}^{1-2\lambda+\alpha}.\end{aligned}$$

As a short computation gives $1 - 2\lambda + \alpha = -P_2(\lambda)/(\lambda(\lambda^2-\lambda+1)) < 0$, we conclude that the integer $\det(C_{i,i+1}, C_{j,j+1})$ vanishes if i is sufficiently large. ■

COROLLARY 6.3. *Suppose that* $\lambda > \lambda_2$. *Then the complement of* J *in* I *is infinite.*

Proof. If $I \setminus J$ were a finite set, then, by the above proposition, all points $C_{i,i+1}$ with $i \in I$ sufficiently large would belong to the same one-dimensional subspace of $\mathbb{R}^2$. By Lemma 2.4, this would imply that $L(C_{i,i+1}) \gg 1$, against the estimates of Lemma 6.1 since $\alpha < 2\lambda - 1 \le 0$. ■

7. Proof of the Theorem. We may assume that $\lambda > \lambda_2 \cong 0.4241 > \sqrt{2}-1$. Then, by Corollary 6.3, there exist infinitely many triples of elements $g < i < j$ of I with i and j consecutive satisfying

$$W_g + W_{g+1} = W_i + W_{i+1} \neq W_j + W_{j+1}. \tag{15}$$

Fix such a triple. Since i and j are consecutive elements of I, we have $W_{i+1} = W_j$ and so

$$W_j = (W_i + W_{i+1}) \cap (W_j + W_{j+1}) = (W_g + W_{g+1}) \cap (W_j + W_{j+1}).$$

Since the sum of $W_g + W_{g+1}$ and $W_j + W_{j+1}$ is the whole of $\mathbb{R}^4$ and since $H(\mathbb{R}^4) = 1$, an application of (3) gives

$$H(W_j) \ll H(W_g + W_{g+1}) H(W_j + W_{j+1}). \tag{16}$$

By Lemma 3.2, we have

$$H(W_g + W_{g+1}) \ll H(W_g)^{-1/\theta} H(W_{g+1}),$$
$$H(W_j + W_{j+1}) \ll H(W_j)^{-1/\theta} H(W_{j+1}),$$

while the estimates (7) of Lemma 5.1 provide

$$H(W_{g+1}) \ll H(W_g)^{\theta} \quad \text{and} \quad H(W_{j+1}) \ll H(W_j)^{\theta}.$$

Using the latter relations respectively as a lower bound for $H(W_g)$ and as an upper bound for $H(W_{j+1})$ and substituting them into the former, we obtain

$$H(W_g+W_{g+1}) \ll H(W_{g+1})^{1-1/\theta^2}, \quad H(W_j+W_{j+1}) \ll H(W_j)^{\theta-1/\theta}. \tag{17}$$

Since $g < i$, we have $X_{g+1} \le X_i$ and so Lemma 3.1 gives

$$H(W_{g+1}) \ll c X_{g+1}^{1-\lambda} \le c X_i^{1-\lambda}. \tag{18}$$

We also have

$$X_i^{1-\lambda} \ll H(W_j)^{\theta^2-1} \tag{19}$$

by the estimates (11) of Corollary 5.3. Combining (16)–(19), we find

$$H(W_j) \ll c^{1-1/\theta^2} H(W_j)^{(1-1/\theta^2)(\theta^2-1)+(\theta-1/\theta)}. \tag{20}$$

Since (19) shows that $H(W_j)$ tends to infinity with i, we conclude that

$$(\theta - 1/\theta)^2 + (\theta - 1/\theta) \ge 1,$$

and so $\theta - 1/\theta \ge 1/\gamma$ where $\gamma = (1+\sqrt{5})/2$ (because $\theta - 1/\theta$ is ≥ 0 and we have $1/\gamma^2 + 1/\gamma = 1$). After simplifications, the latter relation implies

$$\lambda^2 - (1+2\gamma)\lambda + \gamma \ge 0.$$

Since the polynomial $T^2 - (1+2\gamma)T + \gamma$ admits two positive real roots, $\lambda_3 \cong 0.4245$ and $\gamma/\lambda_3 \cong 3.811$, it follows that $\lambda \le \lambda_3$. Moreover, if $\lambda = \lambda_3$, then (20) gives $c \gg 1$, as announced. ∎

Acknowledgments. Part of this work was done during the workshop on Diophantine approximation at the Lorentz Center in Summer 2003. The author thanks the organizers for their invitation and Michel Laurent for several discussions on the topic of the present paper.

References

[1] Y. Bugeaud, *Approximation by Algebraic Numbers*, Cambridge Tracts in Math. 160, Cambridge Univ. Press, 2004.

[2] H. Davenport and W. M. Schmidt, *Approximation to real numbers by quadratic irrationals*, Acta Arith. 13 (1967), 169–176.

[3] —, —, *Approximation to real numbers by algebraic integers*, ibid. 15 (1969), 393–416.

[4] M. Laurent, *Simultaneous rational approximation to the successive powers of a real number*, Indag. Math. (N.S.) 11 (2003), 45–53.

[5] D. Roy, *Approximation simultanée d'un nombre et de son carré*, C. R. Acad. Sci. Paris 336 (2003), 1–6.

[6] —, *Approximation to real numbers by cubic algebraic integers I*, Proc. London Math. Soc. 88 (2004), 42–62.

[7] —, *Approximation to real numbers by cubic algebraic integers II*, Ann. of Math. 158 (2003), 1081–1087.

[8] W. M. Schmidt, *On heights of algebraic subspaces and diophantine approximations*, ibid. 85 (1967), 430–472.

[9] —, *Diophantine Approximations and Diophantine Equations*, Lecture Notes in Math. 1467, Springer, 1991.

Département de Mathématiques
Université d'Ottawa
585 King Edward
Ottawa, ON, Canada K1N 6N5
E-mail: droy@uottawa.ca

Received on 27.12.2007
and in revised form on 8.3.2008 (5600)