ANNALES

DE

L'INSTITUT FOURIER

Damien ROY

Rational approximation to real points on conics

Tome 63, nº 6 (2013), p. 2331-2348.

<http://aif.cedram.org/item?id=AIF_2013__63_6_2331_0>

Ann. Inst. Fourier, Grenoble
63, 6 (2013) 2331-2348

RATIONAL APPROXIMATION TO REAL POINTS ON CONICS

by Damien ROY (*)

ABSTRACT. — A point (ξ_1, ξ_2) with coordinates in a subfield of $\mathbb{R}$ of transcendence degree one over $\mathbb{Q}$, with $1, \xi_1, \xi_2$ linearly independent over $\mathbb{Q}$, may have a uniform exponent of approximation by elements of $\mathbb{Q}^2$ that is strictly larger than the lower bound 1/2 given by Dirichlet's box principle. This appeared as a surprise, in connection to work of Davenport and Schmidt, for points of the parabola $\{(\xi, \xi^2)\,;\, \xi \in \mathbb{R}\}$. The goal of this paper is to show that this phenomenon extends to all real conics defined over $\mathbb{Q}$, and that the largest exponent of approximation achieved by points of these curves satisfying the above condition of linear independence is always the same, independently of the curve, namely $1/\gamma \cong 0.618$ where γ denotes the golden ratio.

RÉSUMÉ. — Un point (ξ_1, ξ_2) à coordonnées dans un sous-corps de $\mathbb{R}$ de degré de transcendance un sur $\mathbb{Q}$, avec $1, \xi_1, \xi_2$ linéairement indépendants sur $\mathbb{Q}$, peut admettre un exposant d'approximation uniforme par les éléments de $\mathbb{Q}^2$ qui soit strictement plus grand que la borne inférieure 1/2 que garantit le principe des tiroirs de Dirichlet. Ce fait inattendu est apparu, en lien avec des travaux de Davenport et Schmidt, pour les points de la parabole $\{(\xi, \xi^2)\,;\, \xi \in \mathbb{R}\}$. Le but de cet article est de montrer que ce phénomène s'étend à toutes les coniques réelles définies sur $\mathbb{Q}$ et que le plus grand exposant d'approximation atteint par les points de ces courbes, sujets à la condition d'indépendance linéaire mentionnée plus tôt, est toujours le même, indépendamment de la courbe, à savoir $1/\gamma \cong 0.618$ où γ désigne le nombre d'or.

1. Introduction

Let n be a positive integer and let $\underline{\xi} = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$. The *uniform exponent of approximation to* $\underline{\xi}$ *by rational points*, denoted $\lambda(\underline{\xi})$, is defined as the supremum of all real numbers λ for which the system of inequalities

$$(1.1) \qquad |x_0| \leqslant X, \quad \max_{1 \leqslant i \leqslant n} |x_0 \xi_i - x_i| \leqslant X^{-\lambda}$$

Keywords: algebraic curves, conics, real points, approximation by rational points, exponent of approximation, simultaneous approximation.
Math. classification: 11J13, 14H50.
(*) Research partially supported by NSERC.

admits a non-zero solution $\mathbf{x} = (x_0, x_1, \dots, x_n) \in \mathbb{Z}^{n+1}$ for each sufficiently large real number $X > 1$. It is one of the classical ways of measuring how well $\underline{\xi}$ can be approximated by elements of $\mathbb{Q}^n$, because each solution of (1.1) with $x_0 \neq 0$ provides a rational point $\mathbf{r} = (x_1/x_0, \dots, x_n/x_0)$ with denominator dividing x_0 such that $\|\underline{\xi} - \mathbf{r}\| \leqslant |x_0|^{-\lambda-1}$, where the symbol $\|\ \|$ stands for the maximum norm. We call it a "uniform exponent" following the terminology of Y. Bugeaud and M. Laurent in [2, §1] because we require a solution of (1.1) for each sufficiently large X (but note that our notation is slightly different as they denote it $\hat{\lambda}(\underline{\xi})$). This exponent depends only on the $\mathbb{Q}$-vector subspace of $\mathbb{R}$ spanned by $1, \xi_1, \dots, \xi_n$ and so, by a result of Dirichlet [12, Chapter II, Theorem 1A], it satisfies $\lambda(\underline{\xi}) \geqslant 1/(s-1)$ where $s \geqslant 1$ denotes the dimension of that subspace. In particular we have $\lambda(\underline{\xi}) = \infty$ when $\underline{\xi} \in \mathbb{Q}^n$, while it is easily shown that $\lambda(\underline{\xi}) \leqslant 1$ when $\underline{\xi} \notin \mathbb{Q}^n$ (see for example [2, Prop. 2.1]).

In their seminal work [3], H. Davenport and W. M. Schmidt determine an upper bound λ_n, depending only on n, for $\lambda(\xi, \xi^2, \dots, \xi^n)$ where ξ runs through all real numbers such that $1, \xi, \dots, \xi^n$ are linearly independent over $\mathbb{Q}$, a condition which amounts to asking that ξ is not algebraic over $\mathbb{Q}$ of degree n or less. Using geometry of numbers, they deduce from this a result of approximation to such ξ by algebraic integers of degree at most $n+1$. In particular they prove that $\lambda(\xi, \xi^2) \leqslant \lambda_2 := 1/\gamma \cong 0.618$ for each non-quadratic irrational real number ξ, where $\gamma = (1+\sqrt{5})/2$ denotes the golden ratio. It is shown in [7, 9] that this upper bound is best possible and, in [8], that the corresponding result of approximation by algebraic integers of degree at most 3 is also best possible. For $n \geqslant 3$, no optimal value is known for λ_n. At present the best known upper bounds are $\lambda_3 \leqslant (1 + 2\gamma - \sqrt{1 + 4\gamma^2})/2 \cong 0.4245$ (see [11]) and $\lambda_n \leqslant 1/\lceil n/2 \rceil$ for $n \geqslant 4$ (see [5]).

As a matter of approaching this problem from a different angle, we propose to extend it to the following setting.

Definition 1.1. — *Let $\mathcal{C}$ be a closed algebraic subset of $\mathbb{R}^n$ of dimension 1 defined over $\mathbb{Q}$, irreducible over $\mathbb{Q}$, and not contained in any proper affine linear subspace of $\mathbb{R}^n$ defined over $\mathbb{Q}$. Then, we put $\lambda(\mathcal{C}) = \sup\{\lambda(\underline{\xi})\,;\, \underline{\xi} \in \mathcal{C}^{li}\}$ where $\mathcal{C}^{li}$ denotes the set of points $\underline{\xi} = (\xi_1, \dots, \xi_n) \in \mathcal{C}$ such that $1, \xi_1, \dots, \xi_n$ are linearly independent over $\mathbb{Q}$.*

Equivalently, such a curve may be described as the Zariski closure over $\mathbb{Q}$ in $\mathbb{R}^n$ of a point $\underline{\xi} \in \mathbb{R}^n$ whose coordinates $\xi_1, \dots, \xi_n$ together with 1 are linearly independent over $\mathbb{Q}$ and generate over $\mathbb{Q}$ a subfield of $\mathbb{R}$

of transcendence degree one. In particular $\mathcal{C}^{li}$ is not empty as it contains that point. From the point of view of metrical number theory the situation is simple since, for the relative Lebesgue measure, almost all points $\underline{\xi}$ of $\mathcal{C}$ have $\lambda(\underline{\xi}) = 1/n$ (see [4]). Of special interest is the curve $\mathcal{C}_n := \{(\xi, \xi^2, \dots, \xi^n) ; \xi \in \mathbb{R}\}$ for any $n \geqslant 2$. As mentioned above, we have $\lambda(\mathcal{C}_2) = 1/\gamma$ and the problem remains to compute $\lambda(\mathcal{C}_n)$ for $n \geqslant 3$. In this paper, we look at the case of conics in $\mathbb{R}^2$ and prove the following result.

THEOREM 1.2. — *Let $\mathcal{C}$ be a closed algebraic subset of $\mathbb{R}^2$ of dimension 1 and degree 2. Suppose that $\mathcal{C}$ is defined over $\mathbb{Q}$ and irreducible over $\mathbb{Q}$. Then, we have $\lambda(\mathcal{C}) = 1/\gamma$. Moreover, the set of points $\underline{\xi} \in \mathcal{C}^{li}$ with $\lambda(\underline{\xi}) = 1/\gamma$ is countably infinite.*

Here the *degree* of $\mathcal{C}$ simply refers to the degree of the irreducible polynomial of $\mathbb{Q}[x_1, x_2]$ defining it. The curve $\mathcal{C}_2$ is the parabola of equation $x_2 - x_1^2 = 0$ but, as we will see, other curves are easier to deal with, for example the curve defined by $x_1^2 - 2 = 0$ which consists of the pair of vertical lines $\{\pm\sqrt{2}\} \times \mathbb{R}$. Note that, for the latter curve, Theorem 1.2 simply says that any $\xi \in \mathbb{R} \setminus \mathbb{Q}(\sqrt{2})$ has $\lambda(\sqrt{2}, \xi) \leqslant 1/\gamma$, with equality defining a denumerable subset of $\mathbb{R} \setminus \mathbb{Q}(\sqrt{2})$. Our main result in the next section provides a slightly finer result.

In [6], it is shown that the cubic $\mathcal{C}$ defined by $x_2 - x_1^3 = 0$ has $\lambda(\mathcal{C}) \leqslant 2(9 + \sqrt{11})/35 \cong 0.7038$, but the case of the line $\sqrt[3]{2} \times \mathbb{R}$ should be simpler to solve and could give ideas to determine the precise value of $\lambda(\mathcal{C})$ for that cubic $\mathcal{C}$. Similarly, looking at lines $(\omega_2, \dots, \omega_n) \times \mathbb{R}$ where $(1, \omega_2, \dots, \omega_n)$ is a basis over $\mathbb{Q}$ of a number field of degree n could provide new ideas to compute $\lambda(\mathcal{C}_n)$.

This paper is organized as follows. In the next section, we state a slightly stronger result in projective setting and note that, for curves $\mathcal{C}$ which are irreducible over $\mathbb{R}$ and contain at least one rational point, the proof simply reduces to the known case of the parabola $\mathcal{C}_2$. In Section 3, we prove the inequality $\lambda(\mathcal{C}) \leqslant 1/\gamma$ for the remaining curves $\mathcal{C}$ by an adaptation of the original argument of Davenport and Schmidt in [3, § 3]. However, the fact that these curves have at most one rational point brings a notable simplification in the proof. In Section 4, we adapt the arguments of [9, § 5] to establish a certain rigidity property for the sequence of minimal points attached to points $\underline{\xi} \in \mathcal{C}^{li}$ with $\lambda(\underline{\xi}) = 1/\gamma$, and deduce from it that the set of these points $\underline{\xi}$ is at most countable. We conclude in Section 5, with the most delicate part, namely the existence of infinitely many points $\underline{\xi} \in \mathcal{C}^{li}$ having exponent $1/\gamma$.

2. The main result in projective framework

For each $n \geqslant 2$, we endow $\mathbb{R}^n$ with the maximum norm, and identify its exterior square $\bigwedge^2 \mathbb{R}^n$ with $\mathbb{R}^{n(n-1)/2}$ via an ordering of the Plücker coordinates. In particular, when $n = 3$, we define the wedge product of two vectors in $\mathbb{R}^3$ as their usual cross-product. We first introduce finer notions of Diophantine approximation in the projective context.

Let $\Xi \in \mathbb{P}^n(\mathbb{R})$ and let $\underline{\Xi} = (\xi_0, \dots, \xi_n)$ be a representative of Ξ in $\mathbb{R}^{n+1}$. We say that a real number $\lambda \geqslant 0$ is an *exponent of approximation* to Ξ if there exists a constant $c = c_1(\underline{\Xi})$ such that the conditions

$$\|\mathbf{x}\| \leqslant X \quad \text{and} \quad \|\mathbf{x} \wedge \underline{\Xi}\| \leqslant cX^{-\lambda}$$

admit a non-zero solution $\mathbf{x} \in \mathbb{Z}^{n+1}$ for each sufficiently large real number X. We say that λ is a *strict exponent of approximation* to Ξ if moreover there exists a constant $c = c_2(\underline{\Xi}) > 0$ such that the same conditions admit no non-zero solution $\mathbf{x} \in \mathbb{Z}^{n+1}$ for arbitrarily large values of X. Both properties are independent of the choice of the representative $\underline{\Xi}$, and we define $\lambda(\Xi)$ as the supremum of all exponents of approximations to Ξ. Clearly, when λ is a strict exponent of approximation to Ξ, we have $\lambda(\Xi) = \lambda$.

Let $T\colon \mathbb{Q}^{n+1} \to \mathbb{Q}^{n+1}$ be an invertible $\mathbb{Q}$-linear map. It extends uniquely to a $\mathbb{R}$-linear automorphism of $\mathbb{R}^{n+1}$ and then to an automorphism of $\mathbb{P}^n(\mathbb{R})$. This defines an action of $\mathrm{GL}_{n+1}(\mathbb{Q})$ on $\mathbb{P}^n(\mathbb{R})$. Moreover, upon choosing an integer $m \geqslant 1$ such that $mT(\mathbb{Z}^{n+1}) \subseteq \mathbb{Z}^{n+1}$, any non-zero point $\mathbf{x} \in \mathbb{Z}^{n+1}$ gives rise to a non-zero point $\mathbf{y} = mT(\mathbf{x}) \in \mathbb{Z}^{n+1}$ satisfying

$$\|\mathbf{y}\| \leqslant c_T\|\mathbf{x}\| \quad \text{and} \quad \|\mathbf{y} \wedge T(\underline{\Xi})\| \leqslant c_T\|\mathbf{x} \wedge \underline{\Xi}\|$$

for a constant $c_T > 0$ depending only on T. Combined with the above definitions, this yields the following invariance property.

Lemma 2.1. — *Let $\Xi \in \mathbb{P}^n(\mathbb{R})$ and $T \in \mathrm{GL}_{n+1}(\mathbb{Q})$. Then we have $\lambda(\Xi) = \lambda(T(\Xi))$. More precisely a real number $\lambda \geqslant 0$ is an exponent of approximation to Ξ, respectively a strict exponent of approximation to Ξ, if and only if it is an exponent of approximation to $T(\Xi)$, respectively a strict exponent of approximation to $T(\Xi)$.*

We also have a natural embedding of $\mathbb{R}^n$ into $\mathbb{P}^n(\mathbb{R})$, sending a point $\underline{\xi} = (\xi_1, \dots, \xi_n)$ to $(1 : \underline{\xi}) := (1 : \xi_1 : \cdots : \xi_n)$. Identifying $\mathbb{R}^n$ with its image in $\mathbb{P}^n(\mathbb{R})$, the above notions of exponent of approximation and strict exponent of approximation carry back to points of $\mathbb{R}^n$. The next lemma, whose proof is left to the reader, shows how they translate in this context and shows moreover that $\lambda(\underline{\xi}) = \lambda(1 : \underline{\xi})$, thus leaving no ambiguity as to the value of $\lambda(\underline{\xi})$.

LEMMA 2.2. — *Let* $\underline{\xi} = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$.

(i) *A real number* $\lambda \geqslant 0$ *is an exponent of approximation to* $(1 : \underline{\xi})$ *if and only if there exists a constant* $c = c_1(\underline{\xi})$ *such that the conditions*

$$|x_0| \leqslant X \quad \text{and} \quad \max_{1 \leqslant i \leqslant n} |x_0 \xi_i - x_i| \leqslant c X^{-\lambda}$$

admit a non-zero solution $\mathbf{x} = (x_0, \dots, x_n) \in \mathbb{Z}^{n+1}$ *for each sufficiently large* X.

(ii) *It is a strict exponent of approximation to* $(1 : \underline{\xi})$ *if and only if there also exists a constant* $c = c_2(\underline{\xi}) > 0$ *such that the above conditions admit no non-zero integer solution for arbitrarily large values of* X.

Finally, we have $\lambda(\underline{\xi}) = \lambda(1 : \underline{\xi})$.

Our main result is the following strengthening of Theorem 1.2.

THEOREM 2.3. — *Let* φ *be a homogeneous polynomial of degree 2 in* $\mathbb{Q}[x_0, x_1, x_2]$. *Suppose that* φ *is irreducible over* $\mathbb{Q}$ *and that its set of zeros* $\mathcal{C}$ *in* $\mathbb{P}^2(\mathbb{R})$ *consists of at least two points.*

(i) *For each point* $\Xi \in \mathcal{C}$ *having* $\mathbb{Q}$*-linearly independent homogeneous coordinates, the number* $1/\gamma$ *is at best a strict exponent of approximation to* Ξ*: if it is an exponent of approximation to* Ξ*, it is a strict one.*

(ii) *There are infinitely many points* $\Xi \in \mathcal{C}$ *which have* $\mathbb{Q}$*-linearly independent homogeneous coordinates and for which* $1/\gamma$ *is an exponent of approximation.*

(iii) *There exists a positive* ϵ*, independent of* φ*, such that the set of points* $\Xi \in \mathcal{C}$ *with* $\lambda(\Xi) > 1/\gamma - \epsilon$ *is countable.*

To show that this implies Theorem 1.2, let $\mathcal{C}$ be as in latter statement. Then, the Zariski closure $\bar{\mathcal{C}}$ of $\mathcal{C}$ in $\mathbb{P}^2(\mathbb{R})$ is infinite and is the zero set of an irreducible homogeneous polynomial of degree 2 in $\mathbb{Q}[x_0, x_1, x_2]$. Moreover, $\mathcal{C}^{li}$ identifies with the set of elements of $\bar{\mathcal{C}}$ with $\mathbb{Q}$-linearly independent homogeneous coordinates. So, if we admit the above theorem, then, in view of Lemma 2.2, Part (i) implies that $\lambda(\mathcal{C}) \leqslant 1/\gamma$, Part (ii) shows that there are infinitely many $\underline{\xi} \in \mathcal{C}^{li}$ with $\lambda(\underline{\xi}) = 1/\gamma$, and Part (iii) shows that the set of points $\underline{\xi} \in \mathcal{C}$ with $\lambda(\underline{\xi}) > 1/\gamma - \epsilon$ is countable. Altogether, this proves Theorem 1.2.

The proof of Part (iii) in Section 4 will show that one can take $\epsilon = 0.005$ but the optimal value for ϵ is probably much larger. In connection to (iii), we also note that the set of elements of $\mathcal{C}$ with $\mathbb{Q}$-linearly dependent homogeneous coordinates is at most countable because each such point belongs to a proper linear subspace of $\mathbb{P}^2(\mathbb{R})$ defined over $\mathbb{Q}$, there are countably many such subspaces, and each of them meets $\mathcal{C}$ in at most two points. So, in order to prove (iii), we may restrict to the points of $\mathcal{C}$ with $\mathbb{Q}$-linearly independent homogeneous coordinates.

Lemma 2.1 implies that, if Theorem 2.3 holds true for a form φ, then it also holds for $\mu(\varphi\circ T)$ for any $T\in \mathrm{GL}_3(\mathbb{Q})$ and any $\mu\in\mathbb{Q}^*$. Thus the next lemma reduces the proof of the theorem to forms of special types.

Lemma 2.4. — *Let φ be an irreducible homogeneous polynomial of $\mathbb{Q}[x_0,x_1,x_2]$ of degree 2 which admits at least two zeros in $\mathbb{P}^2(\mathbb{R})$.*

(i) *If φ is irreducible over $\mathbb{R}$ and admits at least one zero in $\mathbb{P}^2(\mathbb{Q})$, then there exist $\mu\in\mathbb{Q}^*$ and $T\in\mathrm{GL}_3(\mathbb{Q})$ such that $\mu(\varphi\circ T)(x_0,x_1,x_2)=x_0x_2-x_1^2$.*

(ii) *If φ is not irreducible over $\mathbb{R}$, then it admits exactly one zero in $\mathbb{P}^2(\mathbb{Q})$ and there exist $\mu\in\mathbb{Q}^*$ and $T\in\mathrm{GL}_3(\mathbb{Q})$ such that we have $\mu(\varphi\circ T)(x_0,x_1,x_2)=x_0^2-bx_1^2$ for some square-free integer $b>1$.*

(iii) *If φ has no zero in $\mathbb{P}^2(\mathbb{Q})$, then there exist $\mu\in\mathbb{Q}^*$ and $T\in\mathrm{GL}_3(\mathbb{Q})$ such that $\mu(\varphi\circ T)(x_0,x_1,x_2)=x_0^2-bx_1^2-cx_2^2$ for some square-free integers $b>1$ and $c>1$.*

Proof. — We view $(\mathbb{Q}^3,\varphi)$ as a quadratic space. We denote by K its kernel, and by Φ the unique symmetric bilinear form such that $\Phi(\mathbf{x},\mathbf{x})=2\varphi(\mathbf{x})$.

Suppose first that $K\neq\{0\}$. Then, by a change of variables over $\mathbb{Q}$, we can bring φ to a diagonal form $rx_0^2+sx_1^2$ with $r,s\in\mathbb{Q}$. We have $rs\neq 0$ since φ is irreducible over $\mathbb{Q}$, and furthermore $rs<0$ since otherwise the point $(0:0:1)$ would be the only zero of φ in $\mathbb{P}^2(\mathbb{R})$. Thus, φ is not irreducible over $\mathbb{R}$, and $\dim_{\mathbb{Q}} K=1$.

In the case (i), the above observation shows that $\mathbb{Q}^3$ is non-degenerate. Then, since φ has a zero in $\mathbb{P}^2(\mathbb{Q})$, the space $\mathbb{Q}^3$ decomposes as the orthogonal direct sum of a hyperbolic plane H and a non-degenerate line P. We choose bases $\{\mathbf{v}_0,\mathbf{v}_2\}$ for H and $\{\mathbf{v}_1\}$ for P such that $\varphi(\mathbf{v}_0)=\varphi(\mathbf{v}_2)=0$ and $\Phi(\mathbf{v}_0,\mathbf{v}_2)=-\varphi(\mathbf{v}_1)$. Then $\mu=-1/\varphi(\mathbf{v}_1)$ and the linear map $T\colon\mathbb{Q}^3\to\mathbb{Q}^3$ sending the canonical basis of $\mathbb{Q}^3$ to $(\mathbf{v}_0,\mathbf{v}_1,\mathbf{v}_2)$ have the property stated in (i).

In the case (iii), we have $K=\{0\}$ and so we can write $\mathbb{Q}^3$ as an orthogonal direct sum of one-dimensional non-degenerate subspaces P_0, P_1 and P_2. We order them so that the non-zero values of φ on P_0 have opposite sign to those on P_1 and P_2. This is possible since φ is indefinite. Let $\{\mathbf{v}_0\}$ be a basis of P_0 and put $\mu=1/\varphi(\mathbf{v}_0)$. For $i=1,2$, we can choose a basis $\{\mathbf{v}_i\}$ of P_i such that $\mu\varphi(\mathbf{v}_i)$ is a square-free integer. Then μ and the linear map

$T\colon \mathbb{Q}^3 \to \mathbb{Q}^3$ sending the canonical basis of $\mathbb{Q}^3$ to $(\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2)$ have the property stated in (iii).

In the case (ii), the form φ factors over a quadratic extension $\mathbb{Q}(\sqrt{d})$ of $\mathbb{Q}$ as a product $\varphi(\mathbf{x}) = \rho L(\mathbf{x})\bar{L}(\mathbf{x})$ where L is a linear form, $\bar{L}$ its conjugate over $\mathbb{Q}$, and $\rho \in \mathbb{Q}^*$. As φ is irreducible over $\mathbb{Q}$, the linear forms L and $\bar{L}$ are not multiple of each other. Moreover, for a point $\mathbf{a} \in \mathbb{Q}^3$, we have

$$\varphi(\mathbf{a}) = 0 \iff L(\mathbf{a}) = \bar{L}(\mathbf{a}) = 0 \iff (L + \bar{L})(\mathbf{a}) = \sqrt{d}(L - \bar{L})(\mathbf{a}) = 0.$$

Since $L+\bar{L}$ and $\sqrt{d}(L-\bar{L})$ are linearly independent forms with coefficients in $\mathbb{Q}$, this means that the zero set of φ in $\mathbb{Q}^3$ is a line, and so φ has a unique zero in $\mathbb{P}^2(\mathbb{Q})$. As $\Phi(\mathbf{x}, \mathbf{y}) = \rho L(\mathbf{x})\bar{L}(\mathbf{y}) + \rho\bar{L}(\mathbf{x})L(\mathbf{y})$, this line is contained in the kernel K of φ, and so is equal to K. By an earlier observation, this means that, by a change of variables over $\mathbb{Q}$, we may bring φ to a diagonal form $rx_0^2 + sx_1^2$ with $r, s \in \mathbb{Q}$, $rs < 0$. We may further choose r and s so that $-s/r$ is a square-free integer $b > 0$. Then, the same change of variables brings $r^{-1}\varphi$ to $x_0^2 - bx_1^2$. Finally, we have $b \neq 1$ since φ is irreducible over $\mathbb{Q}$. □

3. Proof of the first part of the main theorem

Let φ and $\mathcal{C}$ be as in the statement of Theorem 2.3. Suppose first that φ is irreducible over $\mathbb{R}$ and that $\mathcal{C}\cap\mathbb{P}^2(\mathbb{Q}) \neq \emptyset$. Then, by Lemma 2.4, there exists $T \in \mathrm{GL}_3(\mathbb{Q})$ such that $T^{-1}(\mathcal{C})$ is the zero-set in $\mathbb{P}^2(\mathbb{R})$ of the polynomial $x_0x_2 - x_1^2$. Let Ξ be a point of $\mathcal{C}$ with $\mathbb{Q}$-linearly independent homogeneous coordinates. Its image $T^{-1}(\Xi)$ has homogeneous coordinates $(1 : \xi : \xi^2)$, for some irrational non-quadratic $\xi \in \mathbb{R}$. Then, by [3, Theorem 1a], the number $1/\gamma$ is at best a strict exponent of approximation to $T^{-1}(\Xi)$, and, by Lemma 2.1, the same applies to Ξ. This proves Part (i) of the theorem in that case.

Otherwise, Lemma 2.4 shows that φ has at most one zero in $\mathbb{P}^2(\mathbb{Q})$. Taking advantage of the major simplification that this entails, we proceed as Davenport and Schmidt in [3, §3]. We fix a point $\Xi \in \mathcal{C}$ with $\mathbb{Q}$-linearly independent homogeneous coordinates $(1 : \xi_1 : \xi_2)$ and an exponent of approximation $\lambda \geqslant 1/2$ for Ξ. Then, by Lemma 2.2, there exists a constant $c > 0$ such that, for each sufficiently large X, the system

$$(3.1) \qquad |x_0| \leqslant X, \quad L(\mathbf{x}) := \max\{|x_0\xi_1 - x_1|, |x_0\xi_2 - x_2|\} \leqslant cX^{-\lambda}$$

has a non-zero solution $\mathbf{x} = (x_0, x_1, x_2) \in \mathbb{Z}^3$. To prove Part (i) of Theorem 2.3, we simply need to show that $\lambda \leqslant 1/\gamma$ and that, when $\lambda = 1/\gamma$, the constant c cannot be chosen arbitrarily small.

To this end, we first note that there exists a sequence of points $(\mathbf{x}_i)_{i\geqslant 1}$ in $\mathbb{Z}^3$ such that

(a) their first coordinates X_i form an increasing sequence $1 \leqslant X_1 < X_2 < X_3 < \cdots$,
(b) the quantities $L_i := L(\mathbf{x}_i)$ form a decreasing sequence $1 > L_1 > L_2 > L_3 > \cdots$,
(c) for each $\mathbf{x} = (x_0, x_1, x_2) \in \mathbb{Z}^3$ and each $i \geqslant 1$ with $|x_0| < X_{i+1}$, we have $L(\mathbf{x}) \geqslant L_i$.

Then, each $\mathbf{x}_i$ is a *primitive* point of $\mathbb{Z}^3$, by which we mean that the gcd of its coordinates is 1. Moreover, the hypothesis that (3.1) has a solution for each large enough X implies that

$$L_i \leqslant c X_{i+1}^{-\lambda} \tag{3.2}$$

for each sufficiently large i, say for all $i \geqslant i_0$. Since φ has at most one zero in $\mathbb{P}^2(\mathbb{Q})$, we may further assume that $\varphi(\mathbf{x}_i) \neq 0$ for each $i \geqslant i_0$. Then, upon normalizing φ so that it has integer coefficients, we conclude that $|\varphi(\mathbf{x}_i)| \geqslant 1$ for the same values of i.

Put $\Xi = (1, \xi_1, \xi_2) \in \mathbb{Q}^3$, and let Φ denote the symmetric bilinear form for which $\Phi(\mathbf{x}, \mathbf{x}) = 2\varphi(\mathbf{x})$. Then, upon writing $\mathbf{x}_i = X_i \Xi + \Delta_i$ and noting that $\varphi(\Xi) = 0$, we find

$$\varphi(\mathbf{x}_i) = X_i \Phi(\Xi, \Delta_i) + \varphi(\Delta_i). \tag{3.3}$$

As $\|\Delta_i\| = L_i$, this yields $|\varphi(\mathbf{x}_i)| \leqslant c_1 X_i L_i$ for a constant $c_1 = c_1(\varphi, \Xi) > 0$. Using (3.2), we conclude that, for each $i \geqslant i_0$, we have $1 \leqslant |\varphi(\mathbf{x}_i)| \leqslant c c_1 X_i X_{i+1}^{-\lambda}$, and so

$$X_{i+1}^{\lambda} \leqslant c c_1 X_i. \tag{3.4}$$

We also note that there are infinitely many values of $i > i_0$ for which $\mathbf{x}_{i-1}$, $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$ are linearly independent. For otherwise, all points $\mathbf{x}_i$ with i large enough would lie in a two dimensional subspace V of $\mathbb{R}^3$ defined over $\mathbb{Q}$. As the products $X_i^{-1}\mathbf{x}_i$ converge to Ξ when $i \to \infty$, this would imply that $\Xi \in V$, in contradiction with the hypothesis that Ξ has $\mathbb{Q}$-linearly independent coordinates. Let I denote the set of these indices i.

For $i \in I$, the integer $\det(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})$ is non-zero and [3, Lemma 4] yields

$$1 \leqslant |\det(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1})| \leqslant 6 X_{i+1} L_i L_{i-1} \leqslant 6c^2 X_{i+1}^{1-\lambda} X_i^{-\lambda},$$

thus $X_i^{\lambda} \leqslant 6c^2 X_{i+1}^{1-\lambda}$. Combining this with (3.4), we deduce that $X_i^{\lambda^2} \leqslant (6c^2)^{\lambda} (c c_1 X_i)^{1-\lambda}$ for each $i \in I$, thus $\lambda^2 \leqslant 1 - \lambda$ and so $\lambda \leqslant 1/\gamma$. Moreover, if $\lambda = 1/\gamma$, this yields $1 \leqslant 6c^2(cc_1)^{1/\gamma}$, and so c is bounded below by a positive constant depending only on φ and $\underline{\Xi}$.

4. Proof of the third part of the main theorem

The arguments in [9, §5] can easily be adapted to show that, for some $\epsilon > 0$ there are at most countably many irrational non-quadratic $\xi \in \mathbb{R}$ with $\lambda(1 : \xi : \xi^2) \geqslant 1/\gamma - \epsilon$. This is, originally, an observation of S. Fischler who, in unpublished work, also computed an explicit value for ϵ. The question was later revisited by D. Zelo who showed in [13, Cor. 1.4.7] that one can take $\epsilon = 3.48 \times 10^{-3}$, and who also proved a p-adic analog of this result. More recently, the existence of such ϵ was established by P. Bel, in a larger context where $\mathbb{Q}$ is replaced by a number field K, and $\mathbb{R}$ by a completion of K at some place [1, Theorem 1.3]. By Lemmas 2.1 and 2.4 (i), this proves Theorem 2.3 (iii) when φ is irreducible over $\mathbb{R}$ and has a non-trivial zero in $\mathbb{P}^2(\mathbb{Q})$.

We now consider the complementary case. Using the notation and results of the previous section, we need to show that, when λ is sufficiently close to $1/\gamma$, the point Ξ lies in a countable subset of $\mathcal{C}$. For this purpose, we may assume that $\lambda > 1/2$. The next two lemmas introduce a polynomial $\psi(\mathbf{x}, \mathbf{y})$ with both algebraic and numerical properties analog to that of the operator $[\mathbf{x}, \mathbf{x}, \mathbf{y}]$ from [9, §2] (*cf.* Lemmas 2.1 and 3.1(iii) of [9]).

Lemma 4.1. — *For any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^3$, we define*

$$\psi(\mathbf{x}, \mathbf{y}) := \Phi(\mathbf{x}, \mathbf{y})\mathbf{x} - \varphi(\mathbf{x})\mathbf{y} \in \mathbb{Z}^3.$$

Then, $\mathbf{z} = \psi(\mathbf{x}, \mathbf{y})$ satisfies $\varphi(\mathbf{z}) = \varphi(\mathbf{x})^2\varphi(\mathbf{y})$ and $\psi(\mathbf{x}, \mathbf{z}) = \varphi(\mathbf{x})^2\mathbf{y}$.

Proof. — For any $a, b \in \mathbb{Q}$, we have $\varphi(a\mathbf{x} + b\mathbf{y}) = a^2\varphi(\mathbf{x}) + ab\Phi(\mathbf{x}, \mathbf{y}) + b^2\varphi(\mathbf{y})$. Substituting $a = \Phi(\mathbf{x}, \mathbf{y})$ and $b = -\varphi(\mathbf{x})$ in this equality yields $\varphi(\mathbf{z}) = \varphi(\mathbf{x})^2\varphi(\mathbf{y})$. The formula for $\psi(\mathbf{x}, \mathbf{z})$ follows from the linearity of ψ in its second argument. □

Lemma 4.2. — *Let $i, j \in \mathbb{Z}$ with $i_0 \leqslant i < j$. Then, the point $\mathbf{w} = \psi(\mathbf{x}_i, \mathbf{x}_j) \in \mathbb{Z}^3$ is non-zero and satisfies*

$$\|\mathbf{w}\| \ll X_i^2 L_j + X_j L_i^2 \quad \text{and} \quad L(\mathbf{w}) \ll X_j L_i^2.$$

Here and for the rest of this section, the implied constants depend only on $\underline{\Xi}$, φ, λ and c.

Proof. — Since $\mathbf{x}_i$ and $\mathbf{x}_j$ are distinct primitive elements of $\mathbb{Z}^3$, they are linearly independent over $\mathbb{Q}$. As $\varphi(\mathbf{x}_i) \neq 0$, this implies that $\mathbf{w} = \Phi(\mathbf{x}_i, \mathbf{x}_j)\mathbf{x}_i - \varphi(\mathbf{x}_i)\mathbf{x}_j \neq 0$. By (3.3), we have

$$\varphi(\mathbf{x}_i) = X_i\Phi(\Xi, \Delta_i) + \mathcal{O}(L_i^2)$$

where $\Delta_i = \mathbf{x}_i - X_i\Xi$. Similarly, for $\Delta_j = \mathbf{x}_j - X_j\Xi$, we find

$$\Phi(\mathbf{x}_i, \mathbf{x}_j) = X_j\Phi(\Xi, \Delta_i) + X_i\Phi(\Xi, \Delta_j) + \Phi(\Delta_i, \Delta_j) = X_j\Phi(\Xi, \Delta_i) + \mathcal{O}(X_iL_j).$$

Substituting these expressions in the formula for $\mathbf{w} = \psi(\mathbf{x}_i, \mathbf{x}_j)$, we obtain

$$\begin{aligned} \mathbf{w} &= \big(X_j\Phi(\Xi, \Delta_i) + \mathcal{O}(X_iL_j)\big)(X_i\Xi + \Delta_i) \\ &\quad - \big(X_i\Phi(\Xi, \Delta_i) + \mathcal{O}(L_i^2)\big)(X_j\Xi + \Delta_j) \\ &= \mathcal{O}(X_i^2L_j + X_jL_i^2)\Xi + \mathcal{O}(X_jL_i^2), \end{aligned}$$

and the conclusion follows. □

We will also need the following result, where the set I (defined in Section 3) is endowed with its natural ordering as a subset of $\mathbb{N}$.

Lemma 4.3. — *For each triple of consecutive elements $i < j < k$ in I, the points $\mathbf{x}_i$, $\mathbf{x}_j$ and $\mathbf{x}_k$ are linearly independent. We have*

$$X_j^\alpha \ll X_i \ll X_j^\theta \text{ and } L_i \ll X_j^{-\alpha} \text{ where } \alpha = \frac{2\lambda - 1}{1 - \lambda} \text{ and } \theta = \frac{1 - \lambda}{\lambda}.$$

Proof. — The fact that i and j are consecutive elements of I implies that $\mathbf{x}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_j$ belong to the same 2-dimensional subspace $V_i = \langle \mathbf{x}_i, \mathbf{x}_{i+1} \rangle_\mathbb{R}$ of $\mathbb{R}^3$. Similarly, $\mathbf{x}_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_k$ belong to $V_j = \langle \mathbf{x}_j, \mathbf{x}_{j+1} \rangle_\mathbb{R}$. Thus $\mathbf{x}_i$, $\mathbf{x}_j$ and $\mathbf{x}_k$ span $V_i + V_j = \langle \mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{x}_{j+1} \rangle_\mathbb{R} = \mathbb{R}^3$, and so they are linearly independent. Then, the normal vectors $\mathbf{x}_i \wedge \mathbf{x}_{i+1}$ to V_i and $\mathbf{x}_j \wedge \mathbf{x}_{j+1}$ to V_j are non-parallel and both orthogonal to $\mathbf{x}_j$. So, their cross-product is a non-zero multiple of $\mathbf{x}_j$. Since $\mathbf{x}_j$ is a primitive point of $\mathbb{Z}^3$ and since these normal vectors have integer coordinates, their cross-product is more precisely a non-zero integer multiple of $\mathbf{x}_j$. This yields

$$\begin{aligned} X_j \leqslant \|\mathbf{x}_j\| \ll \|\mathbf{x}_i \wedge \mathbf{x}_{i+1}\| \, \|\mathbf{x}_j \wedge \mathbf{x}_{j+1}\| &\ll (X_{i+1}L_i)(X_{j+1}L_j) \\ &\ll (X_{i+1}X_{j+1})^{1-\lambda}. \end{aligned}$$

If we use the trivial upper bounds $X_{i+1} \leqslant X_j$ and $X_{j+1} \leqslant X_k$ to eliminate X_{i+1} and X_{j+1} from the above estimate, we obtain $X_j \ll X_k^\theta$. If instead we use the upper bounds $X_{i+1} \ll X_i^{1/\lambda}$ and $X_{j+1} \ll X_j^{1/\lambda}$ coming from (3.4), we find instead $X_j^\alpha \ll X_i$. Finally, if we only eliminate X_{j+1} using $X_{j+1} \ll X_j^{1/\lambda}$, we obtain $X_j^{\alpha/\lambda} \ll X_{i+1}$ and thus $L_i \ll X_{i+1}^{-\lambda} \ll X_j^{-\alpha}$. □

Proposition 4.4. — *Suppose that $\lambda \geqslant 0.613$. For each integer $k \geqslant 1$, put $\mathbf{y}_k = \mathbf{x}_{i_k}$ where i_k is the k-th element of I. Then, for each sufficiently large k, the point $\mathbf{y}_{k+1}$ is a rational multiple of $\psi(\mathbf{y}_k, \mathbf{y}_{k-2})$.*

Proof. — For each integer $k \geqslant 1$, let Y_k denote the first coordinate of $\mathbf{y}_k$. Then, according to Lemma 4.3, we have $Y_{k+1}^{\alpha} \ll Y_k \ll Y_{k+1}^{\theta}$ and $L(\mathbf{y}_k) \ll Y_{k+1}^{-\alpha}$, with $\alpha \geqslant 0.5839$ and $\theta \leqslant 0.6314$. Put $\mathbf{w}_k = \psi(\mathbf{y}_k, \mathbf{y}_{k+1})$. By Lemma 4.2, the point $\mathbf{w}_k$ is non-zero, and the above estimates yield

$$L(\mathbf{w}_k) \ll Y_{k+1} L(\mathbf{y}_k)^2 \ll Y_{k+1}^{1-2\alpha} \quad \text{and} \quad \|\mathbf{w}_k\| \ll Y_k^2 L(\mathbf{y}_{k+1}) \ll Y_{k+2}^{-\alpha} Y_k^2$$

(we dropped the term $Y_{k+1}L(\mathbf{y}_k)^2$ in the upper bound for $\|\mathbf{w}_k\|$ because it tends to 0 as $k \to \infty$ while $\|\mathbf{w}_k\| \geqslant 1$). Using these estimates, we find

$$\begin{aligned}
|\det(\mathbf{y}_{k-2}, \mathbf{y}_{k-1}, \mathbf{w}_k)| &\ll \|\mathbf{w}_k\| L(\mathbf{y}_{k-2}) L(\mathbf{y}_{k-1}) + \|\mathbf{y}_{k-1}\| L(\mathbf{y}_{k-2}) L(\mathbf{w}_k) \\
&\ll Y_{k+2}^{-\alpha} Y_k^{2-\alpha^2-\alpha} + Y_{k-1}^{1-\alpha} Y_{k+1}^{1-2\alpha}, \\
&\ll Y_{k+2}^{-\alpha+\theta^2(2-\alpha^2-\alpha)} + Y_{k+1}^{\theta^2(1-\alpha)+1-2\alpha}, \\
|\det(\mathbf{y}_{k-3}, \mathbf{y}_{k-2}, \mathbf{w}_k)| &\ll \|\mathbf{w}_k\| L(\mathbf{y}_{k-3}) L(\mathbf{y}_{k-2}) + \|\mathbf{y}_{k-2}\| L(\mathbf{y}_{k-3}) L(\mathbf{w}_k) \\
&\ll Y_{k+2}^{-\alpha} Y_k^{2-\alpha^3-\alpha^2} + Y_{k-2}^{1-\alpha} Y_{k+1}^{1-2\alpha}, \\
&\ll Y_{k+2}^{-\alpha+\theta^2(2-\alpha^3-\alpha^2)} + Y_{k+1}^{\theta^3(1-\alpha)+1-2\alpha}.
\end{aligned}$$

Thus both determinants tend to 0 as $k \to \infty$ and so, for each sufficiently large k, they vanish. Since, by Lemma 4.3, $\mathbf{y}_{k-3}, \mathbf{y}_{k-2}, \mathbf{y}_{k-1}$ are linearly independent, this implies that, for those k, the point $\mathbf{w}_k$ is a rational multiple of $\mathbf{y}_{k-2}$. As Lemma 4.1 gives $\psi(\mathbf{y}_k, \mathbf{w}_k) = \varphi(\mathbf{y}_k)^2 \mathbf{y}_{k+1}$, we conclude that $\mathbf{y}_{k+1}$ is a rational multiple of $\psi(\mathbf{y}_k, \mathbf{y}_{k-2})$ for each large enough k. $\square$

We end this section with two corollaries. The first one gathers properties of the sequence $(\mathbf{y}_k)_{k\geqslant 1}$ when $\lambda = 1/\gamma$. The second completes the proof of Theorem 2.3(iii).

Corollary 4.5. — *Suppose that $\lambda = 1/\gamma$. Then, the sequence $(\mathbf{y}_k)_{k\geqslant 1}$ consists of primitive points of $\mathbb{Z}^3$ such that $\psi(\mathbf{y}_k, \mathbf{y}_{k-2})$ is an integer multiple of $\mathbf{y}_{k+1}$ for each sufficiently large k. Any three consecutive points of this sequence are linearly independent and, for each $k \geqslant 1$, we have $\|\mathbf{y}_{k+1}\| \asymp \|\mathbf{y}_k\|^{\gamma}$, $L(\mathbf{y}_k) \asymp \|\mathbf{y}_k\|^{-1}$ and $|\varphi(\mathbf{y}_k)| \asymp 1$.*

Proof. — The first assertion simply adds a precision on Proposition 4.4 based on the fact that $\mathbf{y}_{k+1}$ is a primitive integer point. Aside from the estimate for $|\varphi(\mathbf{y}_k)|$, the second assertion is a direct consequence of Lemma 4.3 since, for $\lambda = 1/\gamma$, we have $\alpha = \theta = 1/\gamma$. To complete the proof, we use the estimate $|\varphi(\mathbf{x}_i)| \ll X_i L_i$ established in the previous section as a consequence of (3.3). Since $\varphi(\mathbf{y}_k)$ is a non-zero integer, it yields $1 \leqslant |\varphi(\mathbf{y}_k)| \ll 1$. □

Corollary 4.6. — *Suppose that $\lambda \geqslant 0.613$. Then, Ξ belongs to a countable subset of $\mathcal{C}$.*

Proof. — Since each $\mathbf{y}_k$ is a primitive point of $\mathbb{Z}^3$ with positive first coordinate, the proposition shows that the sequence $(\mathbf{y}_k)_{k\geqslant 1}$ is uniquely determined by its first terms. As there are countably many finite sequences of elements of $\mathbb{Z}^3$ and as the image of $(\mathbf{y}_k)_{k\geqslant 1}$ in $\mathbb{P}^2(\mathbb{R})$ converges to Ξ, the point Ξ belongs to a countable subset of $\mathcal{C}$. □

5. Proof of the second part of the main theorem

By [9, Theorem 1.1], there exist countably many irrational non-quadratic real numbers ξ for which $1/\gamma$ is an exponent of approximation to $(1:\xi:\xi^2)$. Thus Part (ii) of Theorem 2.3 holds for $\varphi = x_0x_2 - x_1^2$ and consequently, by Lemmas 2.1 and 2.4, it holds for any quadratic form $\varphi \in \mathbb{Q}[x_0, x_1, x_2]$ which is irreducible over $\mathbb{R}$ and admits at least one zero in $\mathbb{P}^2(\mathbb{Q})$. These lemmas also show that, in order to complete the proof of Theorem 2.3(ii), we may restrict to a diagonal form $\varphi = x_0^2 - bx_1^2 - cx_2^2$ where $b > 1$ is a square free integer and where c is either 0 or a square-free integer with $c > 1$. In fact, this even covers the case of $\varphi = x_0x_2 - x_1^2$ since $(x_0 + x_1 + x_2)(x_0 - x_1 - x_2) - (x_1 - x_2)^2 = x_0^2 - 2x_1^2 - 2x_2^2$.

We first establish four lemmas which apply to any quadratic form $\varphi \in \mathbb{Q}[x_0, x_1, x_2]$ and its associated symmetric bilinear form Φ with $\Phi(\mathbf{x}, \mathbf{x}) = 2\varphi(\mathbf{x})$. Our first goal is to construct sequences $(\mathbf{y}_i)$ as in Corollary 4.5. On the algebraic side, we first make the following observation.

Lemma 5.1. — *Suppose that $\mathbf{y}_{-1}, \mathbf{y}_0, \mathbf{y}_1 \in \mathbb{Z}^3$ satisfy $\varphi(\mathbf{y}_i) = 1$ for $i = -1, 0, 1$. We extend this triple to a sequence $(\mathbf{y}_i)_{i\geqslant -1}$ in $\mathbb{Z}^3$ by defining recursively $\mathbf{y}_{i+1} = \psi(\mathbf{y}_i, \mathbf{y}_{i-2})$ for each $i \geqslant 1$. We also define $t_i = \Phi(\mathbf{y}_{i+1}, \mathbf{y}_i) \in \mathbb{Z}$ for each $i \geqslant -1$. Then, for any integer $i \geqslant 1$, we have*

(a) $\varphi(\mathbf{y}_{i-2}) = 1$,
(b) $\det(\mathbf{y}_i, \mathbf{y}_{i-1}, \mathbf{y}_{i-2}) = (-1)^{i-1} \det(\mathbf{y}_1, \mathbf{y}_0, \mathbf{y}_{-1})$,
(c) $t_i = \Phi(\mathbf{y}_{i+1}, \mathbf{y}_i) = \Phi(\mathbf{y}_i, \mathbf{y}_{i-2})$,
(d) $\mathbf{y}_{i+1} = t_i\mathbf{y}_i - \mathbf{y}_{i-2}$,
(e) $t_{i+1} = t_i t_{i-1} - t_{i-2}$.

In particular, $t_{-1} = \Phi(\mathbf{y}_0, \mathbf{y}_{-1})$, $t_0 = \Phi(\mathbf{y}_1, \mathbf{y}_0)$ and $t_1 = \Phi(\mathbf{y}_1, \mathbf{y}_{-1})$.

Proof. — By Lemma 4.1, we have $\varphi(\mathbf{y}_{i+1}) = \varphi(\mathbf{y}_i)^2\varphi(\mathbf{y}_{i-2})$ for each $i \geqslant 1$. This yields (a) by recurrence on i. Then, by definition of ψ, the recurrence formula for $\mathbf{y}_{i+1}$ simplifies to

$$\mathbf{y}_{i+1} = \Phi(\mathbf{y}_i, \mathbf{y}_{i-2})\mathbf{y}_i - \mathbf{y}_{i-2} \quad (i \geqslant 1), \tag{5.1}$$

and so $\det(\mathbf{y}_{i+1}, \mathbf{y}_i, \mathbf{y}_{i-1}) = -\det(\mathbf{y}_i, \mathbf{y}_{i-1}, \mathbf{y}_{i-2})$ for each $i \geqslant 1$, by multilinearity of the determinant. This proves (b) by recurrence on i. From (5.1), we deduce that

$$t_i = \Phi(\mathbf{y}_{i+1}, \mathbf{y}_i) = \Phi(\mathbf{y}_i, \mathbf{y}_{i-2})\Phi(\mathbf{y}_i, \mathbf{y}_i) - \Phi(\mathbf{y}_{i-2}, \mathbf{y}_i) = \Phi(\mathbf{y}_i, \mathbf{y}_{i-2}) \; (i \geqslant 1),$$

which is (c). Then (d) is just a rewriting of (5.1). Combining (c) and (d), we find

$$t_{i+1} = \Phi(\mathbf{y}_{i+1}, \mathbf{y}_{i-1}) = t_i\Phi(\mathbf{y}_i, \mathbf{y}_{i-1}) - \Phi(\mathbf{y}_{i-2}, \mathbf{y}_{i-1}) = t_i t_{i-1} - t_{i-2} \; (i \geqslant 1),$$

which is (e). Finally, the formulas given for t_{-1} and t_0 are taken from the definition while the one for t_1 follows from (c). □

The next lemma provides mild conditions under which the norm of $\mathbf{y}_i$ grows as expected.

Lemma 5.2. — *With the notation of the previous lemma, suppose that $1 \leqslant t_{-1} < t_0 < t_1$ and that $1 \leqslant \|\mathbf{y}_{-1}\| < \|\mathbf{y}_0\| < \|\mathbf{y}_1\|$. Then, $(t_i)_{i\geqslant -1}$ and $(\|\mathbf{y}_i\|)_{i\geqslant -1}$ are strictly increasing sequences of positive integers with $t_{i+1} \asymp t_i^\gamma$ and $\|\mathbf{y}_{i+1}\| \asymp t_{i+2} \asymp \|\mathbf{y}_i\|^\gamma$.*

Here and below, the implied constants are simply meant to be independent of i.

Proof. — Lemma 5.1(e) implies, by recurrence on i, that the sequence $(t_i)_{i\geqslant -1}$ is strictly increasing and, more precisely, that it satisfies

$$(t_i - 1)t_{i-1} < t_{i+1} < t_i t_{i-1} \quad (i \geqslant 1), \tag{5.2}$$

which by [10, Lemma 5.2] implies that $t_{i+1} \asymp t_i^\gamma$. In turn, Lemma 5.1(d) implies, by recurrence on i, that the sequence $(\|\mathbf{y}_i\|)_{i\geqslant -1}$ is strictly increasing with

$$(t_i - 1)\|\mathbf{y}_i\| < \|\mathbf{y}_{i+1}\| < (t_i + 1)\|\mathbf{y}_i\| \quad (i \geqslant 1). \tag{5.3}$$

Combining this with (5.2), we find that the ratios $\rho_i = \|\mathbf{y}_i\|/t_{i+1}$ satisfy

$$(1 - 1/t_i)\rho_i \leqslant \rho_{i+1} \leqslant \frac{1 + 1/t_i}{1 - 1/t_{i+1}}\rho_i \leqslant \frac{1}{(1 - 1/t_i)^2}\rho_i \quad (i \geqslant 1),$$

and so $\rho_1 c_1 \leqslant \rho_i \leqslant \rho_1/c_1^2$ for each $i \geqslant 1$ where $c_1 = \prod_{i\geqslant 1}(1 - 1/t_i) > 0$ is a converging infinite product because t_i tends to infinity with i faster than any geometric series. This means that $\rho_i \asymp 1$, thus $\|\mathbf{y}_i\| \asymp t_{i+1}$, and so $\|\mathbf{y}_{i+1}\| \asymp t_{i+2} \asymp \|\mathbf{y}_i\|^\gamma$ because $t_{i+2} \asymp t_{i+1}^\gamma$. □

For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, we denote by $\langle \mathbf{x}, \mathbf{y}\rangle$ their standard scalar product. When $\mathbf{x} \neq 0$ and $\mathbf{y} \neq 0$, we also denote by $[\mathbf{x}]$, $[\mathbf{y}]$ their respective classes in $\mathbb{P}^2(\mathbb{R})$, and define the *projective distance* between these classes by

$$\operatorname{dist}([\mathbf{x}], [\mathbf{y}]) = \frac{\|\mathbf{x} \wedge \mathbf{y}\|}{\|\mathbf{x}\| \, \|\mathbf{y}\|}.$$

It is not strictly speaking a distance on $\mathbb{P}^2(\mathbb{R})$ but it behaves almost like a distance since it satisfies

$$\operatorname{dist}([\mathbf{x}], [\mathbf{z}]) \leqslant \operatorname{dist}([\mathbf{x}], [\mathbf{y}]) + 2 \operatorname{dist}([\mathbf{y}], [\mathbf{z}])$$

for any non-zero $\mathbf{z} \in \mathbb{R}^3$ (see [10, § 2]). Moreover, the open balls for the projective distance form a basis of the usual topology on $\mathbb{P}^2(\mathbb{R})$. We can now prove the following result.

Lemma 5.3. — *With the notation and hypotheses of Lemmas 5.1 and 5.2, suppose that* $\mathbf{y}_{-1}, \mathbf{y}_0$ *and* $\mathbf{y}_1$ *are linearly independent. Then there exists a zero* $\Xi = (1, \xi_1, \xi_2)$ *of* φ *in* $\mathbb{R}^3$ *with* $\mathbb{Q}$*-linearly independent coordinates such that* $\|\Xi \wedge \mathbf{y}_i\| \asymp \|\mathbf{y}_i\|^{-1}$ *for each* $i \geqslant 1$*. Moreover,* $1/\gamma$ *is an exponent of approximation to the corresponding point* $\Xi = (1 : \xi_1 : \xi_2) \in \mathbb{P}^2(\mathbb{R})$*.*

Proof. — Our first goal is to show that $([\mathbf{y}_i])_{i\geqslant 1}$ is a Cauchy sequence in $\mathbb{P}^2(\mathbb{R})$ with respect to the projective distance. To this end, we use freely the estimates of the previous lemma and define $\mathbf{z}_i = \mathbf{y}_i \wedge \mathbf{y}_{i+1}$ for each $i \geqslant 1$. By Lemma 5.1(b), the points $\mathbf{y}_{i-1}$, $\mathbf{y}_i$ and $\mathbf{y}_{i+1}$ are linearly independent for each $i \geqslant 0$. Thus, none of the products $\mathbf{z}_i$ vanish, and so their norm is at least 1. Moreover, Lemma 5.1(d) applied first to $\mathbf{y}_{i+1}$ and then to $\mathbf{y}_i$ yields

$$\mathbf{z}_i = \mathbf{y}_{i-2} \wedge \mathbf{y}_i = t_{i-1}\mathbf{y}_{i-2} \wedge \mathbf{y}_{i-1} - \mathbf{y}_{i-2} \wedge \mathbf{y}_{i-3} = t_{i-1}\mathbf{z}_{i-2} + \mathbf{z}_{i-3}. \tag{5.4}$$

The above equality $\mathbf{z}_i = \mathbf{y}_{i-2} \wedge \mathbf{y}_i$ with i replaced by $i-3$ implies that

$$\|\mathbf{z}_{i-3}\| \leqslant 2\|\mathbf{y}_{i-5}\| \, \|\mathbf{y}_{i-3}\| \ll t_{i-4}t_{i-2} \asymp t_{i-1}t_{i-5}^{-1} \leqslant t_{i-1}t_{i-5}^{-1}\|\mathbf{z}_{i-2}\|.$$

In view of (5.4), this means that $\|\mathbf{z}_i\| = t_{i-1}(1 + \mathcal{O}(t_{i-5}^{-1})) \, \|\mathbf{z}_{i-2}\|$, and thus

$$\frac{\|\mathbf{z}_i\|}{t_i} = \frac{t_{i-1}t_{i-2}}{t_i}(1 + \mathcal{O}(t_{i-5}^{-1}))\frac{\|\mathbf{z}_{i-2}\|}{t_{i-2}} = (1 + \mathcal{O}(t_{i-5}^{-1}))\frac{\|\mathbf{z}_{i-2}\|}{t_{i-2}}$$

since, by Lemma 5.1(e), we have $t_{i-1}t_{i-2} = t_i(1+t_{i-3}t_i^{-1}) = t_i(1+O(t_{i-5}^{-1}))$. As the series $\sum_{i\geqslant 1} t_i^{-1}$ converges, the same is true of the infinite products

$\prod_{i\geqslant i_0}(1+ct_i^{-1})$ for any $c\in\mathbb{R}$. Thus the above estimates implies that $\|\mathbf{z}_i\|\asymp t_i$, and so we find

$$\mathrm{dist}([\mathbf{y}_i],[\mathbf{y}_{i+1}])=\frac{\|\mathbf{z}_i\|}{\|\mathbf{y}_i\|\,\|\mathbf{y}_{i+1}\|}\asymp\frac{t_i}{t_{i+1}t_{i+2}}\asymp t_{i+1}^{-2}\asymp\|\mathbf{y}_i\|^{-2}.$$

As the series $\sum_{i\geqslant 1}2^i t_{i+1}^{-2}$ is convergent, we deduce that $([\mathbf{y}_i])_{i\geqslant 1}$ forms a Cauchy sequence in $\mathbb{P}^2(\mathbb{R})$, and that its limit $\Xi\in\mathbb{P}^2(\mathbb{R})$ satisfies $\mathrm{dist}([\mathbf{y}_i],\Xi)\asymp\|\mathbf{y}_i\|^{-2}$. In terms of a representative $\underline{\Xi}$ of Ξ in $\mathbb{R}^3$, this means that

$$\|\mathbf{y}_i\wedge\underline{\Xi}\|\asymp\|\mathbf{y}_i\|^{-1}. \tag{5.5}$$

To prove that $\underline{\Xi}$ has $\mathbb{Q}$-linearly independent coordinates, we use the fact that

$$\|\langle\mathbf{u},\mathbf{y}_i\rangle\underline{\Xi}-\langle\mathbf{u},\underline{\Xi}\rangle\mathbf{y}_i\|\leqslant 2\|\mathbf{u}\|\,\|\mathbf{y}_i\wedge\underline{\Xi}\|$$

for any $\mathbf{u}\in\mathbb{R}^3$ [10, Lemma 2.2]. So, if $\langle\mathbf{u},\underline{\Xi}\rangle=0$ for some $\mathbf{u}\in\mathbb{Z}^3$, then, by (5.5), we obtain $|\langle\mathbf{u},\mathbf{y}_i\rangle|\ll\|\mathbf{y}_i\|^{-1}$ for all i. Then, as $\langle\mathbf{u},\mathbf{y}_i\rangle$ is an integer, it vanishes for each sufficiently large i, and so $\mathbf{u}=0$ because any three consecutive $\mathbf{y}_i$ span $\mathbb{R}^3$. This proves our claim. In particular, the first coordinate of $\underline{\Xi}$ is non-zero, and we may normalize $\underline{\Xi}$ so that it is 1. Then, as i goes to infinity, the points $\|\mathbf{y}_i\|^{-1}\mathbf{y}_i$ converge to $\|\underline{\Xi}\|^{-1}\underline{\Xi}$ in $\mathbb{R}^3$ and, since $\varphi(\|\mathbf{y}_i\|^{-1}\mathbf{y}_i)=\|\mathbf{y}_i\|^{-2}$ tends to 0, we deduce that $\varphi(\underline{\Xi})=0$. Finally, $1/\gamma$ is an exponent of approximation to Ξ because, for each $X\geqslant\|\mathbf{y}_1\|$, there exists an index $i\geqslant 1$ such that $\|\mathbf{y}_i\|\leqslant X\leqslant\|\mathbf{y}_{i+1}\|$ and then, by (5.5), the point $\mathbf{x}:=\mathbf{y}_i$ satisfies both

$$\|\mathbf{x}\|\leqslant X\quad\text{and}\quad\|\mathbf{x}\wedge\underline{\Xi}\|\asymp\|\mathbf{y}_i\|^{-1}\asymp\|\mathbf{y}_{i+1}\|^{-1/\gamma}\leqslant X^{-1/\gamma}.$$

□

The last lemma below will enable us to show that the above process leads to infinitely many limit points Ξ.

Lemma 5.4. — *Suppose that $(\mathbf{y}_i)_{i\geqslant -1}$ and $(\mathbf{y}'_i)_{i\geqslant -1}$ are constructed as in Lemma 5.1 and that both of them satisfy the hypotheses of the three preceding lemmas. Suppose moreover that their images in $\mathbb{P}^2(\mathbb{R})$ have the same limit Ξ. Then there exists an integer a such that $\mathbf{y}'_i=\pm\mathbf{y}_{i+a}$ for each $i\geqslant\max\{-1,-1-a\}$.*

Proof. — Let $\underline{\Xi}=(1,\xi_1,\xi_2)$ be a representative of Ξ in $\mathbb{R}^3$, and for each $\mathbf{x}\in\mathbb{Z}^3$ define $L(\mathbf{x})$ as in (3.1). The estimates of Lemma 5.3 imply that $L(\mathbf{y}_i)\asymp\|\mathbf{y}_i\|^{-1}$ and $L(\mathbf{y}'_i)\asymp\|\mathbf{y}'_i\|^{-1}$. For each sufficiently large index j, we can find an integer $i\geqslant 2$ such that $\|\mathbf{y}_{i-1}\|^{3/2}\leqslant\|\mathbf{y}'_j\|\leqslant\|\mathbf{y}_i\|^{3/2}$ and the standard estimates yield

$$\begin{aligned}|\det(\mathbf{y}_{i-1},\mathbf{y}_i,\mathbf{y}'_j)| &\ll \|\mathbf{y}'_j\|L(\mathbf{y}_i)L(\mathbf{y}_{i-1}) + \|\mathbf{y}_i\|L(\mathbf{y}_{i-1})L(\mathbf{y}'_j)\\ &\ll \|\mathbf{y}_i\|^{3/2}\|\mathbf{y}_i\|^{-1}\|\mathbf{y}_{i-1}\|^{-1} + \|\mathbf{y}_i\|\,\|\mathbf{y}_{i-1}\|^{-1}\|\mathbf{y}_{i-1}\|^{-3/2}\\ &\ll \|\mathbf{y}_i\|^{1/2-1/\gamma} = o(1),\end{aligned}$$

and similarly $|\det(\mathbf{y}_i,\mathbf{y}_{i+1},\mathbf{y}'_j)| \ll \|\mathbf{y}_i\|^{-1/(2\gamma)} = o(1)$. Thus, both determinants vanish when j is large enough and then $\mathbf{y}'_j$ is a rational multiple of $\mathbf{y}_i$. However, both points are primitive elements of $\mathbb{Z}^3$ since φ takes value 1 on each of them. So, we must have $\mathbf{y}'_j = \pm\mathbf{y}_i$. Since the two sequences have the same type of growth, we conclude that there exist integers a and $i_0 \geqslant \max\{-1,-1-a\}$ such that $\mathbf{y}'_i = \pm\mathbf{y}_{i+a}$ for each $i \geqslant i_0$. Choose i_0 smallest with this property. If $i_0 \geqslant \max\{0,-a\}$, then, using Lemma 4.1, we obtain

$$\mathbf{y}'_{i_0-1} = \psi(\mathbf{y}'_{i_0+1},\mathbf{y}'_{i_0+2}) = \psi(\pm\mathbf{y}_{i_0+1+a},\pm\mathbf{y}_{i_0+2+a}) = \pm\mathbf{y}_{i_0-1+a}$$

in contradiction with the choice of i_0. Thus we must have $i_0 = \max\{-1,-1-a\}$. □

In view of the remarks made at the beginning of this section, the last result below completes the proof of Theorem 2.3(ii).

Proposition 5.5. — *Let $b > 1$ be a square-free integer and let c be either 0 or a square-free integer with $c > 1$. Then the quadratic form $\varphi = x_0^2 - bx_1^2 - cx_2^2$ admits infinitely many zeros in $\mathbb{P}^2(\mathbb{R})$ which have $\mathbb{Q}$-linearly independent homogeneous coordinates and for which $1/\gamma$ is an exponent of approximation.*

Proof. — The Pell equation $x_0^2-bx_1^2=1$ admits infinitely many solutions in positive integers. We choose one such solution $(x_0,x_1)=(m,n)$. For the other solutions $(m',n')\in(\mathbb{N}^*)^2$, the quantity $mm'-bnn'$ behaves asymptotically like $m'/(m+n\sqrt{b})$ as $m'\to\infty$ and thus, we have $m < mm'-bnn' < m'$ as soon as m' is large enough. We fix such a solution (m',n'). We also choose a pair of integers $r,t>0$ such that $r^2-ct^2=1$. Then, the three points

$$\mathbf{y}_{-1}=(1,0,0),\quad \mathbf{y}_0=(m,n,0)\quad\text{and}\quad \mathbf{y}_1=(rm',rn',t)$$

are $\mathbb{Q}$-linearly independent. They satisfy

$$\|\mathbf{y}_{-1}\|=1<\|\mathbf{y}_0\|=m<rm'\leqslant\|\mathbf{y}_1\|\quad\text{and}\quad \varphi(\mathbf{y}_i)=1\quad(i=-1,0,1).$$

For such a triple, consider the corresponding sequences $(t_i)_{i\geqslant-1}$ and $(\mathbf{y}_i)_{i\geqslant-1}$ as defined in Lemma 5.1. The symmetric bilinear form attached to φ being $\Phi = 2(x_0y_0 - bx_1y_1 - cx_2y_2)$, we find

$$t_{-1} = 2m < t_0 = 2r(mm' - bnn') < t_1 = 2rm'.$$

Therefore the hypotheses of Lemmas 5.2 and 5.3 are fulfilled and so the sequence $([\mathbf{y}_i])_{i\geqslant -1}$ converges in $\mathbb{P}^2(\mathbb{R})$ to a zero Ξ of φ which has $\mathbb{Q}$-linearly independent homogeneous coordinates and for which $1/\gamma$ is an exponent of approximation. To complete the proof and show that there are infinitely many such points, it suffices to prove that any other choice of m, n, m', n', r, t as above leads to a different limit point. Clearly, it leads to a different sequence $(\mathbf{y}'_i)_{i\geqslant -1}$. If $[\mathbf{y}'_i]$ and $[\mathbf{y}_i]$ converge to the same point Ξ as $i \to \infty$, then by Lemma 5.4, there exists $a \in \mathbb{Z}$ such that $\mathbf{y}'_i = \pm \mathbf{y}_{i+a}$ for each $i \geqslant \max\{-1, -1-a\}$. But, in both sequences $(\mathbf{y}_i)_{i\geqslant -1}$ and $(\mathbf{y}'_i)_{i\geqslant -1}$, the first point is the only one of norm 1, and moreover the first three points have non-negative entries. So, we must have $a = 0$ and $\mathbf{y}'_i = \mathbf{y}_i$ for $i = -1, 0, 1$, a contradiction. □

BIBLIOGRAPHY

[1] P. Bel, “Approximation simultanée d’un nombre v-adique et de son carré par des nombres algébriques”, J. Number Theory, to appear.

[2] Y. Bugeaud & M. Laurent, “Exponents of Diophantine approximation and Sturmian continued fractions”, *Ann. Inst. Fourier (Grenoble)* **55** (2005), no. 3, p. 773-804.

[3] H. Davenport & W. M. Schmidt, “Approximation to real numbers by algebraic integers”, *Acta Arith.* **15** (1968/1969), p. 393-416.

[4] D. Kleinbock, “Extremal subspaces and their submanifolds”, *Geom. Funct. Anal.* **13** (2003), no. 2, p. 437-466.

[5] M. Laurent, “Simultaneous rational approximation to the successive powers of a real number”, *Indag. Math. (N.S.)* **14** (2003), no. 1, p. 45-53.

[6] S. Lozier & D. Roy, “Simultaneous approximation to a real number and to its cube”, Acta Arith., to appear.

[7] D. Roy, “Approximation simultanée d’un nombre et de son carré”, *C. R. Math. Acad. Sci. Paris* **336** (2003), no. 1, p. 1-6.

[8] ———, “Approximation to real numbers by cubic algebraic integers. II”, *Ann. of Math. (2)* **158** (2003), no. 3, p. 1081-1087.

[9] ———, “Approximation to real numbers by cubic algebraic integers. I”, *Proc. London Math. Soc. (3)* **88** (2004), no. 1, p. 42-62.

[10] ———, “On two exponents of approximation related to a real number and its square”, *Canad. J. Math.* **59** (2007), no. 1, p. 211-224.

[11] ———, “On simultaneous rational approximations to a real number, its square, and its cube”, *Acta Arith.* **133** (2008), no. 2, p. 185-197.

[12] W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, vol. 785, Springer, Berlin, 1980, x+299 pages.

[13] D. Zelo, *Simultaneous approximation to real and p-adic numbers*, ProQuest LLC, Ann Arbor, MI, 2009, Thesis (Ph.D.)–University of Ottawa (Canada), 147 pages.

Manuscrit reçu le 27 janvier 2012,
accepté le 3 avril 2012.

Damien ROY
Université d'Ottawa
Département de Mathématiques
585 King Edward
Ottawa, Ontario K1N 6N5 (Canada)
droy@uottawa.ca