Author's personal copy

Monatsh Math (2015) 177:397–419
DOI 10.1007/s00605-015-0769-9

Nesterenko's linear independence criterion for vectors

Stéphane Fischler[1]

Received: 12 December 2013 / Accepted: 15 April 2015 / Published online: 15 May 2015
© Springer-Verlag Wien 2015

Abstract In this paper we deduce a lower bound for the rank of a family of p vectors in $\mathbb{R}^k$ (considered as a vector space over the rationals) from the existence of a sequence of linear forms on $\mathbb{R}^p$, with integer coefficients, which are small at k points. This is a generalization to vectors of Nesterenko's linear independence criterion (which corresponds to $k = 1$), used by Ball–Rivoal to prove that infinitely many values of Riemann zeta function at odd integers are irrational. The proof is based on geometry of numbers, namely Minkowski's theorem on convex bodies.

Keywords Linear independence · Irrationality · Zeta values

Mathematics Subject Classification Primary 11J72; Secondary 11J13 · 11H06

1 Introduction

The motivation for this paper comes from irrationality results on values of Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$ at odd integers $s \geq 3$. The first result is due to Apéry [1]: $\zeta(3) \notin \mathbb{Q}$. The next breakthrough in this topic is due to Rivoal [22] and Ball–Rivoal [2]:

$$\dim_{\mathbb{Q}} \operatorname{Span}_{\mathbb{Q}}(1, \zeta(3), \zeta(5), \zeta(7), \ldots, \zeta(a)) \geq \frac{\log a}{1 + \log 2}(1 + o(1)) \qquad (1.1)$$

Communicated by A. Constantin.

✉ Stéphane Fischler
stephane.fischler@math.u-psud.fr

[1] Équipe d'Arithmétique et de Géométrie Algébrique, Université Paris-Sud,
Bâtiment 425, 91405 Orsay Cedex, France

Author's personal copy

as $a \to \infty$, where a is an odd integer; notice this is a lower bound on the rank of this family of real numbers, in $\mathbb{R}$ considered as a vector space over the rationals. Conjecturally the left handside is equal to $\frac{a+1}{2}$, but even the constant $\frac{1}{1+\log 2}$ in Eq. (1.1) has never been improved. Actually, known refinements of Ball–Rivoal's proof provide sharper lower bounds only for fixed values of a: the improvement always lies inside the error term $o(1)$ as $a \to \infty$.

However, the following improvement of (1.1) is proved in [10]:

Theorem 1 *Let $\varepsilon > 0$, and a be an odd integer sufficiently large with respect to ε. Then letting N denote the integer part of $\frac{1-\varepsilon}{1+\log 2}\log a$, there exist odd integers $\sigma_1, \ldots, \sigma_N$ between 3 and a such that:*

- *1, $\zeta(\sigma_1)$, ..., $\zeta(\sigma_N)$ are linearly independent over the rationals;*
- *For any $i \neq j$, $|\sigma_i - \sigma_j| > a^\varepsilon$.*

In particular, if there are only N odd integers σ between 3 and a such that $\zeta(\sigma)$ is irrational, then they have to be evenly distributed (see [10]).

The strategy for proving Theorem 1 is based on the following classical construction. For non-negative integers β, b, n, r with β and b odd, $1 \leq \beta \leq b$, and $2br < a$, let

$$J_{\beta,n} = \frac{d_{2n}^{a+b-1}(2n)!^{a-2br}}{(\beta-1)!}\sum_{k=1}^{\infty}\frac{\mathrm{d}^{\beta-1}}{\mathrm{d}k^{\beta-1}}\left(\frac{(k-2rn)_{2rn}^{b}(k+2n+1)_{2rn}^{b}}{(k)_{2n+1}^{a}}\right), \tag{1.2}$$

where the derivative is taken at k, Pochhammer's symbol is defined by $(\alpha)_p = \alpha(\alpha+1)\ldots(\alpha+p-1)$, and d_{2n} is the least common multiple of 1, 2, 3, ..., $2n$. It is not difficult to prove that

$$\begin{aligned} J_{\beta,n} = \widetilde{\ell}_{\beta,n} + \ell_{3,n}\binom{\beta+1}{\beta-1}\zeta(\beta+2) + \ell_{5,n}\binom{\beta+3}{\beta-1}\zeta(\beta+4) \\ + \cdots + \ell_{a,n}\binom{\beta+a-2}{\beta-1}\zeta(\beta+a-1) \end{aligned}$$

with integers $\widetilde{\ell}_{\beta,n}$ and $\ell_{i,n}$; moreover $J_{\beta,n}$ tends to 0 as $n \to \infty$, for any β, provided the parameters satisfy suitable relations (and up to technicalities, see [10] for precise statements). This can be seen as a sequence (L_n) of linear forms on $\mathbb{R}^{(a+b)/2}$, with integer coefficients, that take small values $L_n(e_j) = J_{2j-1,n}$ at $k = \frac{b+1}{2}$ points $e_1, \ldots, e_k \in \mathbb{R}^{(a+b)/2}$. The key point in the proof of Theorem 1 is then to apply the following result, to which the present paper is devoted.

We let $\mathbb{R}^p$ be endowed with its canonical scalar product and the corresponding norm.

Theorem 2 *Let $1 \leq k \leq p-1$, and $e_1, \ldots, e_k \in \mathbb{R}^p$.*

Let $\tau_1, \ldots, \tau_k > 0$ be pairwise distinct real numbers.

Let $(Q_n)_{n\geq 1}$ be an increasing sequence of positive integers, such that $Q_{n+1} = Q_n^{1+o(1)}$.

Author's personal copy

For any $n \geq 1$, let $L_n = \ell_{1,n}X_1 + \cdots + \ell_{p,n}X_p$ be a linear form on $\mathbb{R}^p$, with integer coefficients $\ell_{i,n}$ such that, as $n \to \infty$:

$$|L_n(e_j)| = Q_n^{-\tau_j + o(1)} \textit{ for any } j \in \{1, \dots, k\} \textit{ and } \max_{1 \leq i \leq p} |\ell_{i,n}| \leq Q_n^{1+o(1)}.$$

Then:

(i) *If F is a subspace of $\mathbb{R}^p$ defined over $\mathbb{Q}$ which contains e_1, ..., e_k then*

$$\dim F \geq k + \tau_1 + \cdots + \tau_k.$$

In other words, letting $C_1, \dots, C_p \in \mathbb{R}^k$ denote the columns of the matrix whose rows are $e_1, \dots, e_k \in \mathbb{R}^p$, we have

$$\mathrm{rk}_{\mathbb{Q}}(C_1, \dots, C_p) \geq k + \tau_1 + \cdots + \tau_k$$

in $\mathbb{R}^k$ seen as a $\mathbb{Q}$-vector space.

(ii) *The vectors $e_1, \dots, e_k$ are $\mathbb{R}$-linearly independent in $\mathbb{R}^p$, and the $\mathbb{R}$-subspace they span does not intersect $\mathbb{Q}^p \setminus \{(0, \dots, 0)\}$.*

(iii) *Let $\varepsilon > 0$, and Q be sufficiently large (in terms of ε). Let $\mathcal{C}(\varepsilon, Q)$ denote the set of all vectors that can be written as $\lambda_1 e_1 + \cdots + \lambda_k e_k + u$ with:*

$$\begin{cases} \lambda_1, \dots, \lambda_k \in \mathbb{R} \textit{ such that } |\lambda_j| \leq Q^{\tau_j - \varepsilon} \textit{ for any } j \in \{1, \dots, k\} \\ u \in (\mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_k))^{\perp} \textit{ such that } \|u\| \leq Q^{-1-\varepsilon} \end{cases}$$

Then $\mathcal{C}(\varepsilon, Q) \cap \mathbb{Z}^p = \{(0, \dots, 0)\}$.

If $k = 1$ this is exactly Nesterenko's 1985 linear independence criterion [21] used in the proof of Ball–Rivoal's result (1.1).

In the conclusions, (ii) is an easy result, and (iii) is the main part (it is a quantitative version of (ii)). We deduce (i) from (iii) using Minkowski's convex body theorem, thereby generalizing the proof given in [12,13] of Nesterenko's linear independence criterion. The equivalence between both statements of (i) comes from linear algebra; it is proved in §3.1. The proof of (iii) relies on a matrix lemma (see §3.5) which might be of independent interest and provides some information on linear independence of the linear forms.

A result analogous to Theorem 2, but in which p linearly independent linear forms like L_n appear in the assumption, is proved in §4.3. This linear independence criterion (in the style of Siegel's) is much easier to prove than Theorem 2. Both results can be thought of as transference principles. In this respect it is worth pointing out that in Theorem 2 we assume essentially that for any positive integer Q there is a linear form : indeed this is L_n, where n is such that $Q_n \leq Q < Q_{n+1}$ so that $Q = Q_n^{1+o(1)}$ because $Q_{n+1} = Q_n^{1+o(1)}$. The assumptions imply that this linear form belongs to some convex body, and conclusion (iii) asserts that (up to Q^{ε}) the dual convex body does not contain any non-zero integer point. Therefore it is reasonable to imagine that (iii) is an optimal conclusion up to Q^{ε}. In general the lower bound $k + \tau_1 + \cdots + \tau_k$

Author's personal copy

in (i) is optimal too (see [11] for a converse statement, valid almost everywhere). In the special case $p = 2, k = 1$, and $e_1 = (1, \xi)$, Theorem 2 (iii) yields an upper bound $\mu(\xi) \leq 1 + \frac{1}{\tau_1}$ on the irrationality exponent of ξ, and reduces essentially to Lemma 1 of [12]. A converse statement in this case is proved in [12] (Theorem 1).

The assumption that $\tau_1, \ldots, \tau_k$ are pairwise distinct is very important in Theorem 2, and it cannot be omitted. For instance, if $\tau_1 = \tau_2$ then $L_n(e_1 - e_2)$ could be very small: up to replacing (e_1, e_2) with $(e_1 + e_2, e_1 - e_2)$, this amounts to dropping the assumption that the linear forms L_n are not too small at the points e_j. Now this assumption is known to be essential, already in the classical case of Nesterenko's linear independence criterion (except for proving the linear independence of three numbers, see Theorem 2 of [13]). Actually, if $\tau_1 = \tau_2$ then $L_n(e_1 - e_2)$ could even vanish, so the possibility that $e_1 = e_2$ cannot be eliminated: even assertion (ii) may fail to hold.

We shall prove Theorem 2 in a more general form, stated in §2, which allows the sequences $(|L_n(e_j)|)_{n\geq 1}$ to oscillate (as in [9]), and takes into account divisors of the coefficients $\ell_{i,n}$ (as in [13]); the former is used in [10] to prove Theorem 1. We also include a refinement useful when L_n is not too large at some other point, which is new even in the classical case of Nesterenko's linear independence (with $k = 1$).

We hope that our results will have Diophantine applications besides those of [10]; we mention some directions in §4.4, connected to polylogarithms or zeta values. Our criterion could be used also for q-analogues, as in [13].

The structure of this text is as follows. In §2 we state our result in a very general form, of which Theorem 2 is a special case. Section 3 is devoted to the proof; then we deduce some corollaries in §§4.1 and 4.2. We prove an analogous result in the style of Siegel's linear independence criterion in §4.3, and conclude in §4.4 with Diophantine applications.

2 Statement of the criterion

The following generalization of Theorem 2 is our main result.

Theorem 3 *Let $1 \leq k \leq p - 1$, and $e_1, \ldots, e_k \in \mathbb{R}^p$. Let $(v_1, \ldots, v_p)$ denote a basis of $\mathbb{R}^p$. Let $\tau_1, \ldots, \tau_k > 0$, $\sigma_1 \geq \ldots \geq \sigma_p > 0$, $\omega_1, \ldots, \omega_k$, $\varphi_1, \ldots, \varphi_k$ be real numbers, with $\tau_1, \ldots, \tau_k$ pairwise distinct. Assume that there exist infinitely many integers n with the following property: for any $j \in \{1, \ldots, k\}$, $n\omega_j + \varphi_j \not\equiv \frac{\pi}{2} \mod \pi$.*

Let $(Q_n)_{n\geq 1}$ be an increasing sequence of positive integers, such that $Q_{n+1} = Q_n^{1+O(1/n)}$; if $\omega_1 = \cdots = \omega_k = 0$, this assumption can be weakened to $Q_{n+1} = Q_n^{1+o(1)}$.

For any $n \geq 1$, let $L_n = \ell_{1,n}X_1 + \cdots + \ell_{p,n}X_p$ be a linear form on $\mathbb{R}^p$, with integer coefficients $\ell_{i,n}$ such that, as $n \to \infty$:

$$|L_n(e_j)| = Q_n^{-\tau_j + o(1)} |\cos(n\omega_j + \varphi_j) + o(1)| \text{ for any } j \in \{1, \ldots, k\}, \tag{2.1}$$

and

$$|L_n(v_i)| \leq Q_n^{\sigma_i + o(1)} \text{ for any } i \in \{1, \ldots, p\}.$$

Author's personal copy

For all $n \geq 1$ and $i \in \{1, \dots, p\}$, let $\delta_{i,n}$ be a positive divisor of $\ell_{i,n}$ such that:

(i) *$\delta_{i,n}$ divides $\delta_{i+1,n}$ for any $n \geq 1$ and any $i \in \{1, \dots, p-1\}$,*
(ii) *$\frac{\delta_{j,n}}{\delta_{i,n}}$ divides $\frac{\delta_{j,n+1}}{\delta_{i,n+1}}$ for any $n \geq 1$ and any $0 \leq i < j \leq p$, with $\delta_{0,n} = 1$,*
(iii) *$\delta_{i,n} = Q_n^{d_i + o(1)}$ as $n \to \infty$ for any $i \in \{1, \dots, p\}$, with real numbers d_i such that $0 \leq d_1 \leq \dots \leq d_p \leq \sigma_p$.*

Then:

(i) *If F is a subspace of $\mathbb{R}^p$ defined over $\mathbb{Q}$ which contains e_1, ..., e_k, then $s = \dim F$ satisfies $s \geq k+1$ and*

$$\sigma_1 + \dots + \sigma_{s-k} \geq \tau_1 + \dots + \tau_k + d_1 + \dots + d_s. \tag{2.2}$$

In other words, letting $C_1, \dots, C_p \in \mathbb{R}^k$ denote the columns of the matrix whose rows are $e_1, \dots, e_k \in \mathbb{R}^p$, the rank s of the family $(C_1, \dots, C_p)$ in $\mathbb{R}^k$ seen as a $\mathbb{Q}$-vector space satisfies $s \geq k+1$ and Eq. (2.2).
(ii) *The vectors $e_1, \dots, e_k$ are $\mathbb{R}$-linearly independent in $\mathbb{R}^p$, and the $\mathbb{R}$-subspace they span does not intersect $\mathbb{Q}^p \backslash \{(0, \dots, 0)\}$.*
(iii) *Let $\varepsilon > 0$, and Q be sufficiently large (in terms of ε). Let $\mathcal{C}(\varepsilon, Q)$ denote the set of all vectors that can be written as $\lambda_1 e_1 + \dots + \lambda_k e_k + u$ with:*

$$\begin{cases} \lambda_1, \dots, \lambda_k \in \mathbb{R} \text{ such that } |\lambda_j| \leq Q^{\tau_j - \varepsilon} \text{ for any } j \in \{1, \dots, k\} \\ u \in (\mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_k))^{\perp} \text{ such that } u = \mu_1 v_1 + \dots + \mu_p v_p \text{ with } |\mu_i| \leq Q^{-\sigma_i - \varepsilon} \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for any } i \in \{1, \dots, p\}. \end{cases}$$

Let $\Lambda(Q)$ denote the set of all $(x_1, \dots, x_p) \in \mathbb{Q}^p$ such that $\delta_{i,\Psi(Q)} x_i \in \mathbb{Z}$ for any $i \in \{1, \dots, p\}$, where $\Psi(Q)$ is the largest integer n such that $Q_n \leq Q$. Then $\mathcal{C}(\varepsilon, Q) \cap \Lambda(Q) = \{(0, \dots, 0)\}$.

In the special case where $\sigma_i = \delta_{i,n} = 1$, $d_i = \omega_j = \varphi_j = 0$ for any i, j, n, and $(v_1, \dots, v_p)$ is the canonical basis of $\mathbb{R}^p$, this is exactly Theorem 2 stated in the introduction. Indeed Eq. (2.1) reads $|L_n(e_j)| = Q_n^{-\tau_j + o(1)}$ in this case, and we have $L_n(v_i) = \ell_{i,n}$; moreover Eq. (2.2) reads

$$\dim F \geq k + \tau_1 + \dots + \tau_k.$$

There is only a minor difference in (iii), where the norm of u is the Euclidean one in Theorem 2, and the supremum one in Theorem 3; of course this is not significant.

The real numbers ω_j and φ_j allow oscillating behaviors of the sequences $(|L_n(e_j)|)_{n \geq 1}$. This is used in [10], where the saddle point method is applied. In the special case of Theorem 2 with $k = 1$, the corresponding generalization of Nesterenko's linear independence criterion has been proved in [9] when $Q_n = \beta^n$ for some $\beta > 1$ (which is the most interesting case). We generalize it here to any sequence (Q_n) such that $Q_{n+1} = Q_n^{1+O(1/n)}$; eventhough this assumption is slightly more restrictive than the usual one $Q_{n+1} = Q_n^{1+o(1)}$, it is general enough to include sequences $Q_n = \beta^{n^d}$ with $\beta > 1$ and $d > 0$.

Author's personal copy

The divisors $\delta_{i,n}$ allow one to make use of divisibility properties of the coefficients $\ell_{i,n}$: for instance, in most constructions of linear forms in zeta values, $\ell_{i,n}$ is a multiple of $\delta_{i,n} = d_n^{e_i}$ for some $e_i \geq 1$, where $d_n = \operatorname{lcm}(1, 2, \dots, n)$. The first refinement of Nesterenko's linear independence criterion involving such divisors $\delta_{i,n}$ is Theorem 1 of [13], which is essentially the special case of Theorem 3 (i) where $k = 1$, $\sigma_i = 1$, $\omega_j = \varphi_j = 0$, and $(v_1, \dots, v_p)$ is the canonical basis of $\mathbb{R}^p$; it is the main ingredient in the proof [13] that $1, \zeta(3)$ and $\zeta(j)$ are $\mathbb{Q}$-linearly independent for some odd integer j between 5 and 139.

The real numbers σ_i allow one to take advantage of the fact that the linear forms L_n might be smaller than $\|L_n\|$ at some given points v_i (eventhough $L_n(v_i)$ does not tend to 0 as $n \to \infty$). For instance, if $(v_1, \dots, v_p)$ is the canonical basis, this is useful when one has a sharper upper bound on $|\ell_{i,n}|$ for some values of i than for others. This feature is new even in the case of Nesterenko's linear independence criterion (namely, with $k = 1$, $\sigma_i = \delta_{i,n} = 1$, and $d_i = \omega_j = \varphi_j = 0$). It would be interesting to deduce from this refinement a Diophantine consequence. Actually it happens for linear forms in zeta values that $\lim_{n\to\infty} |\ell_{i,n}|^{1/n}$ exists for any i and does depend on i. For instance, F. Amoroso and T. Rivoal have noticed that in the expansion of

$$n!^{a-1} \sum_{k=1}^{\infty} \frac{(k-n)_n}{(k)_{n+1}^a}$$

as a linear combination of zeta values, the coefficients of odd and even zeta values don't have the same size (provided a is even).

It is very important in Theorem 3 that $\tau_1, \dots, \tau_k$ are pairwise distinct; however it is not always necessary to compute their exact values. For instance, if $\min(\tau_1, \dots, \tau_k)$ is greater than or equal to some $\tau > 0$, then Eq. (2.2) implies

$$\sigma_1 + \dots + \sigma_{s-k} \geq k\tau + d_1 + \dots + d_s;$$

in the special case of Theorem 2 this lower bound reads $\dim F \geq k(1 + \tau)$. This remark is already used (with $k = 1$) in [2], and also in the proof [10] of Theorem 1. We refer to §4.2 below for a related result.

At last, notice that if the assumptions of Theorem 3 hold with $e_1, \dots, e_k$, then they hold also if we forget one of the e_j's (say e_k, with $k \geq 2$). The same implication holds also for parts (ii) and (iii) of the conclusion, since the convex body $\mathcal{C}(\varepsilon, Q)$ becomes smaller when e_k is omitted. However this implication does not hold for part (i); to fix this we refine part (i) in the following corollary (which is used in [10]).

Corollary 1 *In the situation of Theorem 3, assume also that $\tau_1 > \dots > \tau_k$. Then for any subspace F of $\mathbb{R}^p$ defined over $\mathbb{Q}$ we have*

$$s \geq t + 1 \text{ and } \sigma_1 + \dots + \sigma_{s-t} \geq \tau_{k+1-t} + \dots + \tau_k + d_1 + \dots + d_s, \tag{2.3}$$

provided that $s = \dim F$ and $t = \dim(F \cap \operatorname{Span}_{\mathbb{R}}(e_1, \dots, e_k))$ are positive.

In other words, for any surjective $\mathbb{R}$-linear map $\pi : \mathbb{R}^k \to \mathbb{R}^t$ with $t \geq 1$, Eq. (2.3) holds with

Author's personal copy

$$s = \mathrm{rk}_{\mathbb{Q}}(\pi(C_1), \dots, \pi(C_p))$$

where the rank is computed in $\mathbb{R}^t$ seen as a $\mathbb{Q}$-vector space.

Proof of Corollary 1 Let F be a subspace of $\mathbb{R}^p$ defined over $\mathbb{Q}$; assume that $s = \dim F$ and $t = \dim(F \cap \mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_k))$ are positive. For any $j \in \{1, \dots, k\}$ we let $D_j = \dim(F \cap \mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_j))$, so that $0 \leq D_1 \leq \dots \leq D_k = t$ and $D_j \in \{D_{j-1}, D_{j-1} + 1\}$ for any j (with $D_0 = 0$). Then there exist t integers $1 \leq j_1 < \dots < j_t \leq k$ such that $D_j = D_{j-1} + 1$ if, and only if, j is among the j_i's. For any $i \in \{1, \dots, t\}$, there exists $e'_i \in F \cap \mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_{j_i})$ such that $e'_i \notin \mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_{j_i-1})$. Then we have $e'_i = \sum_{j=1}^{j_i} \lambda_{i,j} e_j$ for real numbers $\lambda_{i,j}$ such that $\lambda_{i,j_i} \neq 0$. Since $\tau_1 > \dots > \tau_{j_i}$, Eq. (2.1) yields

$$|L_n(e'_i)| = Q_n^{-\tau_{j_i} + o(1)} |\cos(n\omega_{j_i} + \varphi_{j_i}) + o(1)|.$$

Therefore Theorem 3 applies to $e'_1, \dots, e'_t$ with $\tau_{j_1}, \dots, \tau_{j_t}$. Since $\tau_1 > \dots > \tau_k$, the inequality (2.2) obtained in this way implies Eq. (2.3). This concludes the proof of Corollary 1, except for the second part of the conclusion which will be proved at the end of §3.1 below.

3 Proof of the criterion

This section is devoted to proving Theorem 3, of which Theorem 2 stated in the introduction is a special case (see §2). Reindexing $e_1, \dots, e_k$ is necessary, we assume that $\tau_1 > \dots > \tau_k > 0$. This assumption will be used in §§3.3 and 3.6.

3.1 Rational rank of vectors

In this section, we give some details about the conclusions of our criterion, which allow us to prove the equivalence of both conclusions of (i) in Theorems 2 and 3, and to conclude the proof of Corollary 1.

In Nesterenko's linear independence criterion, a lower bound is derived for the dimension of the $\mathbb{Q}$-subspace of $\mathbb{R}$ spanned by $\xi_0, \dots, \xi_r \in \mathbb{R}$, that is, for the $\mathbb{Q}$-rank of $\xi_0, \dots, \xi_r$ in $\mathbb{R}$ considered as a vector space over $\mathbb{Q}$. This rank is equal to the dimension of the smallest subspace of $\mathbb{R}^{r+1}$, defined over the rationals, which contains the point $(\xi_0, \dots, \xi_r)$. We generalize this equality to our setting in Lemma 1 below.

Recall that a subspace F of $\mathbb{R}^p$ is said to be *defined over* $\mathbb{Q}$ if it is the zero locus of a family of linear forms with rational coefficients. This is equivalent to the existence of a basis (or a generating family) of F, as a vector space over $\mathbb{R}$, consisting in vectors of $\mathbb{Q}^p$ (see for instance §8 of [3]). Since the intersection of a family of subspaces of $\mathbb{R}^p$ defined over $\mathbb{Q}$ is again defined over $\mathbb{Q}$, there exists for any subset $S \subset \mathbb{R}^p$ a minimal subspace of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains S: this is the intersection of all subspaces of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contain S.

Let M be a matrix with $k \geq 1$ rows, $p \geq 1$ columns, and real entries. Letting $e_1, \dots, e_k \in \mathbb{R}^p$ denote the rows of M, we can consider as above the smallest subspace

Author's personal copy

of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains $e_1, \dots, e_k$. On the other hand, we denote by $C_1, \dots, C_p \in \mathbb{R}^k$ the columns of M and consider $\mathbb{R}^k$ as an infinite-dimensional vector space over $\mathbb{Q}$. Then $\text{Span}_{\mathbb{Q}}(C_1, \dots, C_p)$ is the smallest $\mathbb{Q}$-vector subspace of $\mathbb{R}^k$ containing $C_1, \dots, C_p$; it consists in all linear combinations $r_1 C_1 + \dots + r_p C_p$ with $r_1, \dots, r_p \in \mathbb{Q}$. Its dimension (as a $\mathbb{Q}$-vector space) is the rank (over $\mathbb{Q}$) of $C_1, \dots, C_p$, denoted by $\text{rk}_{\mathbb{Q}}(C_1, \dots, C_p)$.

Lemma 1 *Let $M \in \text{Mat}_{k,p}(\mathbb{R})$ with $k, p \geq 1$. Denote by $e_1, \dots, e_k \in \mathbb{R}^p$ denote the rows of M, and by $C_1, \dots, C_p \in \mathbb{R}^k$ its columns. Then $\text{rk}_{\mathbb{Q}}(C_1, \dots, C_p)$ is the dimension of the smallest subspace of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains $e_1, \dots, e_k$.*

When $k = 1$, this lemma means that the $\mathbb{Q}$-rank of $\xi_0, \dots, \xi_r$ is equal to the dimension of the smallest subspace of $\mathbb{R}^{r+1}$, defined over the rationals, which contains the point $(\xi_0, \dots, \xi_r)$.

Proof of Lemma 1 Let $G = (\text{Span}_{\mathbb{R}}(e_1, \dots, e_k))^{\perp}$, where $\mathbb{R}^p$ is equipped with the usual scalar product. Let F denote the minimal subspace of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains $e_1, \dots, e_k$. Then $F^{\perp}$ is the maximal subspace of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which is contained in $G = \{e_1, \dots, e_k\}^{\perp}$. Therefore $F^{\perp} = \text{Span}_{\mathbb{R}}(G \cap \mathbb{Q}^p) = (G \cap \mathbb{Q}^p) \otimes_{\mathbb{Q}} \mathbb{R}$: any basis of the $\mathbb{Q}$-vector space $G \cap \mathbb{Q}^p$ is an $\mathbb{R}$-basis of $F^{\perp}$. Since $G \cap \mathbb{Q}^p = \ker \psi$ where $\psi : \mathbb{Q}^p \to \mathbb{R}^k$ is defined by $\psi(r_1, \dots, r_p) = r_1 C_1 + \dots + r_p C_p$, we have:

$$\dim_{\mathbb{R}} F = p - \dim_{\mathbb{R}} F^{\perp} = p - \dim_{\mathbb{Q}}(G \cap \mathbb{Q}^p) = \text{rk}_{\mathbb{Q}} \psi = \text{rk}_{\mathbb{Q}}(C_1, \dots, C_p).$$

This concludes the proof of Lemma 1.

Let us deduce from Lemma 1 the following generalization, and use it to prove the second assertion of Corollary 1.

Lemma 2 *Let $M, e_1, \dots, e_k, C_1, \dots, C_p$ be as in Lemma 1. Let $\pi : \mathbb{R}^k \to \mathbb{R}^t$ be a $\mathbb{R}$-linear map, with $t \geq 1$. Then the rank of $(\pi(C_1), \dots, \pi(C_p))$ in $\mathbb{R}^t$ (seen as a $\mathbb{Q}$-vector space) is equal to the dimension of the minimal subspace F of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains the image of $\psi \circ {}^t\pi$; here ψ is the $\mathbb{R}$-linear map of the dual of $\mathbb{R}^k$ to $\mathbb{R}^p$ which maps the canonical basis to $(e_1, \dots, e_k)$.*

Proof of Lemma 2 Let P be the matrix of π with respect to canonical bases, and $M' = PM$. Applying Lemma 1 to M' gives directly the result.

Proof of the second assertion of Corollary 1 Let F denote the minimal subspace of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains the image of $\psi \circ {}^t\pi$; then Lemma 2 yields $\dim F = s$. Now $\text{rk}({}^t\pi) = \text{rk}(\pi) = t$ and ψ is injective because $e_1, \dots, e_k$ are $\mathbb{R}$-linearly independent (using conclusion (ii) of Theorem 3), so that $\text{Im}(\psi \circ {}^t\pi)$ has dimension t. Since this subspace is contained in both F and $\text{Span}_{\mathbb{R}}(e_1, \dots, e_k) = \text{Im}\psi$, we have $\dim(F \cap \text{Span}_{\mathbb{R}}(e_1, \dots, e_k)) \geq t$. Now the first part of Corollary 1 (deduced in §2 from Theorem 3) shows that Eq. (2.3) holds when t is replaced with this (possibly larger) dimension; therefore it holds with t. This concludes the proof of the second assertion of Corollary 1.

Author's personal copy

3.2 Reduction to the non-oscillatory case

In this subsection, we deduce the general case of Theorem 3 from the special case where $\omega_1 = \cdots = \omega_k = 0$; notice that in this case we have $\phi_j \not\equiv \frac{\pi}{2} \mod \pi$ for any $j \in \{1, \ldots, k\}$, so that Eq. (2.1) reads $|L_n(e_j)| = Q_n^{-\tau_j + o(1)}$. This special case will be proved in the following subsections, under the assumption that $Q_{n+1} = Q_n^{1+o(1)}$ (which is weaker than the assumption $Q_{n+1} = Q_n^{1+O(1/n)}$ we make when $\omega_1, \ldots, \omega_k$ may be non-zero).

Let $\omega_1, \ldots, \omega_k, \varphi_1, \ldots, \varphi_k$, and (Q_n) be as in Theorem 3, with $Q_{n+1} = Q_n^{1+O(1/n)}$. Since there are infinitely many integers n such that, for any $j \in \{1, \ldots, k\}$, $n\omega_j + \varphi_j \not\equiv \frac{\pi}{2} \mod \pi$, Proposition 1 of [9] provides $\varepsilon, \lambda > 0$ and an increasing function $\psi : \mathbb{N} \to \mathbb{N}$ such that $\lim_{n\to\infty} \frac{\psi(n)}{n} = \lambda$ and, for any n and any $j \in \{1, \ldots, k\}$, $|\cos(\psi(n)\omega_j + \varphi_j)| \geq \varepsilon$. Let $L'_n = L_{\psi(n)}$ and $Q'_n = Q_{\psi(n)}$ for any $n \geq 1$. Then we have $|L'_n(e_j)| = {Q'_n}^{-\tau_j + o(1)}$ because $|\cos(\psi(n)\omega_j + \varphi_j)| = Q_{\psi(n)}^{o(1)}$. Let us check that $Q'_{n+1} = Q_n'^{1+o(1)}$; then the special case of Theorem 3 will apply to the sequences $(L'_n)_{n\geq 1}$ and $(Q'_n)_{n\geq 1}$, with the same other parameters: this will conclude the proof.

Since $Q_{n+1} = Q_n^{1+O(1/n)}$ there exists $M > 0$ such that, for any $n \geq 1$, $Q_{n+1} \leq Q_n^{1+M/n}$; this implies

$$\log Q_{n+\ell} \leq (1 + M/n)^\ell \log Q_n$$

for any $\ell \geq 0$. Letting $\delta_n = \psi(n+1) - \psi(n) \geq 1$, we have:

$$\begin{aligned}\log Q'_{n+1} = \log Q_{\psi(n)+\delta_n} \leq (1 + M/\psi(n))^{\delta_n} \log Q_{\psi(n)} &\leq \exp(M\delta_n/\psi(n)) \log Q_{\psi(n)} \\ &= (1 + o(1)) \log Q'_n\end{aligned}$$

since $1 + x \leq e^x$ and $\delta_n = o(n)$ since $\psi(n) = \lambda n + o(n)$. This concludes the reduction to the case where $\omega_1 = \ldots = \omega_k = 0$ and $Q_{n+1} = Q_n^{1+o(1)}$.

3.3 Proof of (ii)

Let us come now to the easiest part of Theorem 3, namely (ii). We shall prove simultaneously that $e_1, \ldots, e_k$ are linearly independent in $\mathbb{R}^p$, and that $F \cap \mathbb{Q}^p = \{(0, \ldots, 0)\}$ where $F = \mathrm{Span}_{\mathbb{R}}(e_1, \ldots, e_k)$. With this aim in mind, we assume (by contradiction) that there exist real numbers $\lambda_1, \ldots, \lambda_k$, not all zero, such that $\sum_{j=1}^k \lambda_j e_j \in \mathbb{Q}^p$; multiplying all λ_j by a common denominator of the coordinates, we may assume $\sum_{j=1}^k \lambda_j e_j \in \mathbb{Z}^p$. Then $\kappa_n = L_n(\sum_{j=1}^k \lambda_j e_j) = \sum_{j=1}^k \lambda_j L_n(e_j)$ is an integer for any $n \geq 1$. Now if n is sufficiently large then $|\kappa_n| \leq \sum_{j=1}^k |\lambda_j| |L_n(e_j)| < 1$, so that $\kappa_n = 0$. Let j_0 denote the largest integer j such that $\lambda_j \neq 0$. Then for any n sufficiently large, the fact that $\kappa_n = 0$ implies $|\lambda_{j_0} L_n(e_{j_0})| = |\sum_{j=1}^{j_0-1} \lambda_j L_n(e_j)|$ so that

Author's personal copy

$$|\lambda_{j_0}| \leq \sum_{j=1}^{j_0-1} |\lambda_j| \frac{|L_n(e_j)|}{|L_n(e_{j_0})|} \leq \sum_{j=1}^{j_0-1} |\lambda_j| Q_n^{\tau_{j_0}-\tau_j+o(1)}$$

as $n \to \infty$. Now the right handside tends to 0 as $n \to \infty$ because we have assumed that $\tau_1 > \cdots > \tau_k$, so that $\lambda_{j_0} = 0$: this contradicts the definition of λ_{j_0}.

Therefore such real numbers $\lambda_1, \ldots, \lambda_k$ cannot exist, and this concludes the proof of (ii).

3.4 Proof that (ii) and (iii) imply (i)

Before proceeding in §§3.5 and 3.6 to the proof of (iii), which is the main part, we deduce (i) from (ii) and (iii). Recall that the second statement of (i) is equivalent to the first one (which we shall prove now) thanks to Lemma 1 proved in §3.1.

Let F be a subspace of $\mathbb{R}^p$, defined over $\mathbb{Q}$, which contains e_1, ..., e_k. Letting $s = \dim F$, we have $s > k$ using (ii). Assertion (iii) yields, for any $\varepsilon > 0$ and any Q sufficiently large (in terms of ε), a subset $\mathcal{C}(\varepsilon, Q)$ and a lattice $\Lambda(Q)$ such that $\mathcal{C}(\varepsilon, Q) \cap \Lambda(Q) = \{(0, \ldots, 0)\}$. Now $\mathcal{C}(\varepsilon, Q) \cap F$ is a convex body, compact and symmetric with respect to the origin, in the Euclidean space F. On the other hand, $\Lambda(Q) \cap F$ is a lattice in F because F is defined over $\mathbb{Q}$. Therefore Minkowski's convex body theorem (see for instance Chapter III of [5]) implies that $\mathcal{C}(\varepsilon, Q) \cap F$ has volume less than $2^s \det(\Lambda(Q) \cap F)$. Letting

$$\alpha = \tau_1 + \cdots + \tau_k - \sigma_1 - \cdots - \sigma_{s-k} - s\varepsilon,$$

this volume is greater than or equal to Q^α, up to a multiplicative constant which depends only on F, $e_1, \ldots, e_k$, v_1, ..., v_p (using the inequalities $\sigma_1 \geq \cdots \geq \sigma_p$). On the other hand, since $d_1 \leq \cdots \leq d_p$ we have $\det(\Lambda(Q) \cap F) \leq c Q^{\beta+o(1)}$ where $\beta = -d_1 - \cdots - d_s$ and c is a constant depending only on F. Since Q can be chosen arbitrarily large, the above-mentioned consequence of Minkowski's theorem yields $\alpha \leq \beta$. Now ε can be any positive real number, so that we obtain

$$\tau_1 + \cdots + \tau_k + d_1 + \cdots + d_s \leq \sigma_1 + \cdots + \sigma_{s-k},$$

thereby concluding the proof of (i).

3.5 A matrix lemma

We state and prove in this section the main tool in the proof of Theorem 2, namely Lemma 3. This result has been used recently by Dauguet [6], and might be of independent interest; its proof relies on estimating the determinant and cofactors.

Lemma 3 *Let A be a $k \times k$ matrix with real positive entries $a_{i,j}$, $1 \leq i, j \leq k$, such that*

Author's personal copy

$$a_{i',j}a_{i,j'} \leq \frac{1}{(k+1)!}a_{i,j}a_{i',j'} \text{ for any } i, j, i', j' \text{ such that } i < i' \text{ and } j < j'. \tag{3.1}$$

Then A is an invertible matrix, and letting $A^{-1} = [b_{i,j}]_{1\leq i,j\leq k}$ we have

$$|b_{j,i}| \leq \Big(1 + \frac{1}{k} + \frac{1}{k^2}\Big)a_{i,j}^{-1} \text{ for any } i, j \in \{1, \ldots, k\}.$$

Lemma 3 is optimal up to the value of the constant $1+\frac{1}{k}+\frac{1}{k^2}$: it would be false with a constant less than $1/k$ instead (this is immediately seen by computing a diagonal coefficient of AA^{-1}, which is equal to 1). We did not try to improve on the constant $1 + \frac{1}{k} + \frac{1}{k^2}$, but anyway it could easily be made smaller by replacing $\frac{1}{(k+1)!}$ in (3.1) with a smaller constant.

In the proof of Lemma 3 we shall use the following result.

Lemma 4 *Under the assumptions of Lemma 3, for any $\sigma \in \mathfrak{S}_k$ we have*

$$\prod_{j=1}^{k} a_{\sigma(j),j} \leq \eta_\sigma \prod_{j=1}^{k} a_{j,j} \tag{3.2}$$

where $\eta_\sigma = \frac{1}{(k+1)!}$ if $\sigma \neq$ Id, and $\eta_{\rm Id} = 1$.

Proof of Lemma 4 For $\sigma \neq$ Id let κ_σ denote the largest integer $j \in \{1, \ldots, k\}$ such that $\sigma(j) \neq j$; put also $\kappa_{\rm Id} = 0$. We are going to prove Eq. (3.2) by induction on κ_σ. If $\kappa_\sigma \leq 1$ then $\sigma =$ Id, so that Eq. (3.2) holds trivially. Let $\sigma \in \mathfrak{S}_k$ be such that $\kappa_\sigma \geq 2$, and assume that Eq. (3.2) holds for any σ' such that $\kappa_{\sigma'} < \kappa_\sigma$. We have $\sigma(j) = j$ for any $j \in \{\kappa_\sigma + 1, \ldots, k\}$, and $\sigma(\kappa_\sigma) < \kappa_\sigma$. Let $j_0 = \sigma^{-1}(\kappa_\sigma)$; then $j_0 < \kappa_\sigma$. Let $\sigma' = \sigma \circ \tau_{j_0,\kappa_\sigma}$ where τ_{j_0,κ_σ} is the transposition that exchanges j_0 and κ_σ. Then $\sigma'(j) = j$ for any $j \in \{\kappa_\sigma, \ldots, k\}$ so that $\kappa_{\sigma'} < \kappa_\sigma$ and Eq. (3.2) holds for σ'. Since $\sigma'(j) = \sigma(j)$ for $j \notin \{j_0, \kappa_\sigma\}$, $\sigma'(j_0) = \sigma(\kappa_\sigma)$ and $\sigma'(\kappa_\sigma) = \kappa_\sigma$, this implies (using the fact that $\eta_{\sigma'} \leq 1$)

$$a_{\sigma(\kappa_\sigma),j_0}a_{\kappa_\sigma,\kappa_\sigma} \prod_{\substack{1\leq j\leq k\\ j\notin\{j_0,\kappa_\sigma\}}} a_{\sigma(j),j} \leq \prod_{j=1}^{k} a_{j,j}.$$

On the other hand, Eq. (3.1) implies

$$a_{\kappa_\sigma,j_0}a_{\sigma(\kappa_\sigma),\kappa_\sigma} \leq \frac{1}{(k+1)!}a_{\sigma(\kappa_\sigma),j_0}a_{\kappa_\sigma,\kappa_\sigma}$$

because $\sigma(\kappa_\sigma) < \kappa_\sigma$ and $j_0 < \kappa_\sigma$. Multiplying out the previous two inequalities yields Eq. (3.2) for σ, since $\sigma(j_0) = \kappa_\sigma$. This concludes the proof of Lemma 4.

Author's personal copy

Proof of Lemma 3 Letting $\Delta = |\det A|$ we have, using Lemma 4:

$$\Delta \geq \prod_{j=1}^{k} a_{j,j} - \sum_{\substack{\sigma\in\mathfrak{S}_k\\ \sigma\neq\mathrm{Id}}} \prod_{j=1}^{k} a_{\sigma(j),j} \geq \Big(1 - \frac{1}{k+1}\Big) \prod_{j=1}^{k} a_{j,j} > 0 \tag{3.3}$$

so that A is invertible. Given $i, j \in \{1, \dots, k\}$ we have $|b_{j,i}| = \frac{\Delta_{i,j}}{\Delta}$ where $\Delta_{i,j}$ is the absolute value of the determinant of the matrix obtained from A by deleting the i-th row and the jth column. Using Lemma 4 again we have

$$\Delta_{i,j} \leq \sum_{\substack{\sigma\in\mathfrak{S}_k\\ \sigma(j)=i}} \prod_{\substack{1\leq j'\leq k\\ j'\neq j}} a_{\sigma(j'),j'} \leq \left(\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \sigma(j)=i}} \eta_\sigma \right) a_{i,j}^{-1} \prod_{j'=1}^{k} a_{j',j'}. \tag{3.4}$$

Now we have $\eta_\sigma = 1$ for at most one σ, and $\eta_\sigma = \frac{1}{(k+1)!}$ for all other permutations σ among the $(k-1)!$ such that $\sigma(j) = i$, so that

$$\sum_{\substack{\sigma\in\mathfrak{S}_k\\ \sigma(j)=i}} \eta_\sigma \leq 1 + \frac{(k-1)!}{(k+1)!} = \frac{k+1+\frac{1}{k}}{k+1}.$$

Combining this upper bound with Eqs. (3.3) and (3.4) yields

$$|b_{j,i}| = \frac{\Delta_{i,j}}{\Delta} \leq \frac{k+1+\frac{1}{k}}{k} a_{i,j}^{-1},$$

thereby completing the proof of Lemma 3.

3.6 Proof of (iii)

We are now in position to prove the remaining part of Theorem 2, namely (iii). We assume $\tau_1 > \cdots > \tau_k > 0$ and $\omega_1 = \cdots = \omega_k = 0$ (see §3.2), so that $|L_n(e_j)| = Q_n^{-\tau_j+o(1)}$.

Before giving details, let us make a few comments on our strategy.

Recall that Nesterenko's linear independence criterion is much easier to prove if the linear forms L_n, L_{n+1}, …, L_{n+p-1} are linearly independent (see §2.3 of [13] or the references to Siegel's criterion in §4.3 below). Of course this is not always the case, but Lemma 3 enables us to make a step in this direction. Actually letting $F = \mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_k)$, we consider the restrictions $L_{n|F}$ of the linear forms to F; recall that $\dim F = k$ thanks to (ii) proved in §3.3. It is not true in general that $L_{n|F}$, $L_{n+1|F}$, …, $L_{n+k-1|F}$ are linearly independent linear forms on F: for instance, the equality $L_n = L_{n+1}$ might hold for any even integer n (because of the error terms $o(1)$ in the assumptions of Theorem 3). To make this statement correct, we introduce

Author's personal copy

a function $\varphi : \mathbb{N}^* \to \mathbb{N}^*$ such that $\varphi(n) \geq n+1$ for any $n \geq 1$. The integer $\varphi(n)$ plays the role of $n+1$, that is: applying φ corresponds to "taking the next integer". The idea is that $\varphi(n)$ will be large enough (in comparison to n) to avoid obvious counter-examples as above coming from error terms. In more precise terms, $\varphi(n)$ will be defined by the property $Q_{\varphi(n)-1} \leq Q_n^{1+\varepsilon_1} < Q_{\varphi(n)}$ (where ε_1 is a small positive real number); in this way, the error terms $o(1)$ in the assumptions of Theorem 3 will not be a problem any more.

With this definition, we shall prove that for any n sufficiently large, the linear forms $L_{n|F}$, $L_{\varphi(n)|F}$, $L_{\varphi_2(n)|F}$, ..., $L_{\varphi_{k-1}(n)|F}$ on F are linearly independent (where $\varphi_i = \varphi \circ \ldots \circ \varphi$), so that they make up a basis of the dual vector space $F^\star$. In the proof of Theorem 3 we shall need the following quantitative version of this property: in writing the linear form $e_j^\star$ (defined by $e_j^\star(\lambda_1 e_1 + \cdots + \lambda_k e_k) = \lambda_j$) as a linear combination of $\frac{1}{L_n(e_j)} L_{n|F}$, $\frac{1}{L_{\varphi(n)}(e_j)} L_{\varphi(n)|F}, \ldots, \frac{1}{L_{\varphi_{k-1}(n)}(e_j)} L_{\varphi_{k-1}(n)|F}$, the coefficients that appear are bounded independently from n (actually they are between -3 and 3): see Eq. (3.8) below. This will follow from Lemma 3 applied to the matrix $A_n = [|L_{\varphi_{i-1}(n)}(e_j)|]_{1\leq i,j\leq k}$. The point in applying this lemma is that sharp upper and lower bounds on $|L_{\varphi_{i-1}(n)}(e_j)|$ are available; the assumption $\tau_1 > \cdots > \tau_k$ plays also a central role here.

Now let us prove (iii).

Let $\varepsilon > 0$. We choose $\varepsilon_1 > 0$ sufficiently small, so that

$$((1+\varepsilon_1)^{k-1} - 1)\max(1, \tau_1, \sigma_1) < \varepsilon/4. \tag{3.5}$$

If $k = 1$ there is no assumption on ε_1, because it does not really appear in the proof: Lemma 3 is a triviality in this case, and the proof of (iii) reduces essentially to that of [13].

For any $n \geq 1$, we define $\varphi(n)$ by $Q_{\varphi(n)-1} \leq Q_n^{1+\varepsilon_1} < Q_{\varphi(n)}$, because the sequence (Q_n) is increasing and we may assume $Q_n \geq 1$ for any n. Then we have $\varphi(n) \geq n+1$. This implies $\lim_{n\to+\infty} \varphi(n) = +\infty$, so that $Q_{\varphi(n)} = Q_{\varphi(n)-1}^{1+o(1)}$ (because we assume $Q_{n+1} = Q_n^{1+o(1)}$) and

$$Q_{\varphi(n)} = Q_n^{1+\varepsilon_1+o(1)}; \tag{3.6}$$

here $o(1)$ denotes any sequence that tends to 0 as $n \to \infty$. Moreover the assumption $|L_n(e_j)| = Q_n^{-\tau_j+o(1)}$ implies $|L_n(e_j)| > 0$ for any j, n with n sufficiently large. We have also for any n sufficiently large and any $j \in \{1, \ldots, k\}$:

$$|L_{\varphi(n)}(e_j)| = Q_{\varphi(n)}^{-\tau_j+o(1)} = Q_n^{-\tau_j(1+\varepsilon_1)+o(1)} < |L_n(e_j)|. \tag{3.7}$$

For $i \in \{0, \ldots, k-1\}$ let $\varphi_i = \varphi \circ \ldots \circ \varphi$ denote the map φ composed i times with itself (so that $\varphi_0(n) = n$ and $\varphi_1(n) = \varphi(n)$). Put

$$A_n = \Big[|L_{\varphi_{i-1}(n)}(e_j)|\Big]_{1\leq i,j\leq k}$$

and denote by $a_{i,j}$ the entries of A_n (omitting for simplicity the dependence on n). Let us check the assumption (3.1) of Lemma 3, provided n is sufficiently large. Let

Author's personal copy

$i, j, i', j' \in \{1, \ldots, k\}$ be such that $i < i'$ and $j < j'$; we put $n' = \varphi_{i-1}(n)$ and $n'' = \varphi_{i'-1}(n)$, so that $n'' \geq \varphi(n')$. Using Eq. (3.6) and the assumption $\tau_j > \tau_{j'}$ we obtain

$$\frac{a_{i',j}a_{i,j'}}{a_{i,j}a_{i',j'}} = \left|\frac{L_{n''}(e_j)L_{n'}(e_{j'})}{L_{n'}(e_j)L_{n''}(e_{j'})}\right| = \frac{Q_{n''}^{\tau_{j'}-\tau_j+o(1)}}{Q_{n'}^{\tau_{j'}-\tau_j+o(1)}} \leq \left(\frac{Q_{\varphi(n')}}{Q_{n'}}\right)^{\tau_{j'}-\tau_j+o(1)}$$
$$= Q_{n'}^{\varepsilon_1(\tau_{j'}-\tau_j)+o(1)} \leq \frac{1}{(k+1)!}$$

if n is sufficiently large, so that Lemma 3 applies. Given $M = \sum_{j=1}^{k} \lambda_j e_j$ with $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, we have $L_{\varphi_{i-1}(n)}(M) = \sum_{j=1}^{k} a_{i,j}\lambda'_j$ where we let $\lambda'_j = \lambda_j$ if $L_{\varphi_{i-1}(n)}(e_j) > 0$, and $\lambda'_j = -\lambda_j$ otherwise. Therefore Lemma 3 yields, for any $j \in \{1, \ldots, k\}$ and any n sufficiently large:

$$|\lambda_j| = |\lambda'_j| = \left|\sum_{i=1}^{k} b_{j,i}L_{\varphi_{i-1}(n)}(M)\right| \leq \left(1 + \frac{1}{k} + \frac{1}{k^2}\right)\sum_{i=1}^{k}\frac{|L_{\varphi_{i-1}(n)}(M)|}{|L_{\varphi_{i-1}(n)}(e_j)|}. \quad (3.8)$$

This upper bound on $|\lambda_j|$ in terms of the $|L_{\varphi_{i-1}(n)}(M)|$ is the main tool we shall use now in the proof.

Let Q be sufficiently large in terms of ε, and assume that $\mathcal{C}(\varepsilon, Q) \cap \Lambda(Q)$ contains a non-zero point P. Then we have

$$P = \lambda_1 e_1 + \cdots + \lambda_k e_k + u = (x_1, \ldots, x_p) \neq (0, \ldots, 0)$$

with $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, $u = \mu_1 v_1 + \cdots + \mu_p v_p \in (\mathrm{Span}_{\mathbb{R}}(e_1, \ldots, e_k))^{\perp}$, $|\lambda_j| \leq Q^{\tau_j-\varepsilon}$ for any $j \in \{1, \ldots, k\}$, $|\mu_i| \leq Q^{-\sigma_i-\varepsilon}$ for any $i \in \{1, \ldots, p\}$, and $\delta_{i,n}x_i \in \mathbb{Z}$ for any i, where $n = \Psi(Q)$ is the largest integer such that $Q_n \leq Q$. In particular we have $Q_n \leq Q < Q_{n+1}$ so that $Q = Q_n^{1+o(1)}$, and n tends to ∞ as $Q \to \infty$: if $u_n = o(1)$, that is $u_n \to 0$ as $n \to \infty$, then u_n tends also to 0 as $Q \to \infty$.

Let ℓ denote the least integer such that

$$\text{for any } j \in \{1, \ldots, k\}, \text{ we have } |\lambda_j L_\ell(e_j)| \leq \frac{\delta_{p,\ell}}{3k\delta_{p,n}}. \quad (3.9)$$

Since $|\lambda_j| \leq Q^{\tau_j-\varepsilon}$ and n is sufficiently large, this upper bound holds for n so that this integer exists and we have $\ell \leq n$.

The integer ℓ depends on Q and on the choice of a non-zero point $P \in \mathcal{C}(\varepsilon, Q) \cap \Lambda(Q)$. Let us prove that $\ell \to \infty$ as $Q \to \infty$, uniformly with respect to the choice of P. Let $\ell_0 \geq 1$, and denote by K_{ℓ_0} the set of all points $P' = \lambda'_1 e_1 + \cdots + \lambda'_k e_k + u'$ with

$$|\lambda'_j| \min_{1 \leq \ell' \leq \ell_0} |L_{\ell'}(e_j)| \leq \frac{1}{3k} \text{ for any } j \in \{1, \ldots, k\},$$

Author's personal copy

where $u' \in (\mathrm{Span}(e_1, \dots, e_k))^\perp$ can be written as $u' = \mu'_1 v_1 + \cdots + \mu'_p v_p$ with $|\mu'_i| \le Q^{-\sigma_i - \varepsilon}$ for any $i \in \{1, \dots, p\}$. By definition of ℓ and K_{ℓ_0}, if $\ell \le \ell_0$ then $\frac{\delta_{p,n}}{\delta_{p,\ell}} P \in K_{\ell_0}$. Moreover the point $\frac{\delta_{p,n}}{\delta_{p,\ell}} P$ belongs also to $\Lambda(Q_{\ell_0})$ since

$$\delta_{i,\ell_0} \left(\frac{\delta_{p,n}}{\delta_{p,\ell}} x_i \right) = \left(\frac{\delta_{i,\ell_0}}{\delta_{i,\ell}} \right) \left(\frac{\delta_{p,n}/\delta_{i,n}}{\delta_{p,\ell}/\delta_{i,\ell}} \right) \left(\delta_{i,n} x_i \right) \in \mathbb{Z}$$

for any $i \in \{1, \dots, p\}$, by assumption on the divisors $\delta_{t,n}$. Therefore (assuming $\ell \le \ell_0$) the point $\frac{\delta_{p,n}}{\delta_{p,\ell}} P$ belongs to $K_{\ell_0} \cap \Lambda(Q_{\ell_0})$, which is a finite set because K_{ℓ_0} is compact and $\Lambda(Q_{\ell_0})$ is discrete. Now the function $\chi : K_{\ell_0} \cap \Lambda(Q_{\ell_0}) \to \mathbb{R}$ defined by $\chi(P') = \|\pi_\perp(P')\|$, where $\pi_\perp$ is the orthogonal projection on $(\mathrm{Span}(e_1, \dots, e_k))^\perp$, has a least positive value χ_0. We have $\chi(\frac{\delta_{p,n}}{\delta_{p,\ell}} P) \ne 0$ because $P \notin \mathbb{Q}^p \cap \mathrm{Span}(e_1, \dots, e_k) = \{(0, \dots, 0)\}$ (using assertion (ii) proved in §3.3), so that

$$\chi_0 \le \chi \left(\frac{\delta_{p,n}}{\delta_{p,\ell}} P \right) = \frac{\delta_{p,n}}{\delta_{p,\ell}} \|u\| \le Q_n^{d_p + o(1)} Q^{-\sigma_p - \varepsilon} = Q^{d_p - \sigma_p - \varepsilon + o(1)}$$

since $\delta_{p,\ell} \ge 1$ and $\sigma_p \le \cdots \le \sigma_1$. This inequality implies that Q is not too large in terms of ℓ_0 and ε (because we assume $d_p \le \sigma_p$). This concludes the proof that $\ell \to \infty$ as $Q \to \infty$. In what follows, a sequence denoted by $o(1)$ will tend to 0 as n, ℓ or Q tends to ∞; therefore in any case, it tends to 0 as $Q \to \infty$. Moreover, we may assume ℓ to be arbitrarily large.

We come back now to the point $P \in \mathcal{C}(\varepsilon, Q) \cap \Lambda(Q)$ chosen above. Since $u = \mu_1 v_1 + \cdots + \mu_p v_p$ with $|\mu_h| \le Q^{-\sigma_h - \varepsilon}$ for any h, we have for any $i \in \{1, \dots, k\}$:

$$\begin{aligned} |L_{\varphi_{i-1}(\ell)}(u)| &\le \sum_{h=1}^{p} |\mu_h| |L_{\varphi_{i-1}(\ell)}(v_h)| \le \sum_{h=1}^{p} Q^{-\sigma_h - \varepsilon} Q_{\varphi_{i-1}(\ell)}^{\sigma_h + o(1)} \\ &\le \sum_{h=1}^{p} Q_n^{-\sigma_h - \varepsilon + o(1)} Q_\ell^{\sigma_h (1+\varepsilon_1)^{i-1} + o(1)} \text{ using Eq. (3.6)} \\ &\le \sum_{h=1}^{p} \left(\frac{Q_\ell}{Q_n} \right)^{\sigma_h} Q_n^{-\varepsilon + o(1)} Q_\ell^{\varepsilon/4 + o(1)} \text{ using Eq. (3.5) and } \sigma_h \le \sigma_1 \\ &\le \left(\frac{Q_\ell}{Q_n} \right)^{d_p} Q^{-\varepsilon/2} < \frac{1}{3} \frac{\delta_{p,\ell}}{\delta_{p,n}} \text{ since } \sigma_h \ge \sigma_p \ge d_p \text{ and } \ell \le n. \end{aligned} \tag{3.10}$$

On the other hand, Eqs. (3.9) and (3.7) yield for any $i \in \{1, \dots, k\}$:

$$\left| L_{\varphi_{i-1}(\ell)} \left(\sum_{j=1}^{k} \lambda_j e_j \right) \right| \le \sum_{j=1}^{k} \frac{|L_{\varphi_{i-1}(\ell)}(e_j)|}{|L_\ell(e_j)|} \frac{\delta_{p,\ell}}{3k\delta_{p,n}} \le \frac{\delta_{p,\ell}}{3\delta_{p,n}},$$

Author's personal copy

since ℓ is sufficiently large. Combining this inequality with Eq. (3.10) we obtain for the point $P = \lambda_1 e_1 + \cdots + \lambda_k e_k + u$:

$$|L_{\varphi_{i-1}(\ell)}(P)| \leq \frac{\delta_{p,\ell}}{3\delta_{p,n}} + \frac{\delta_{p,\ell}}{3\delta_{p,n}} < \frac{\delta_{p,\ell}}{\delta_{p,n}}. \tag{3.11}$$

Now we have $L_{\varphi_{i-1}(\ell)} = \ell_{1,\varphi_{i-1}(\ell)} X_1 + \cdots + \ell_{p,\varphi_{i-1}(\ell)} X_p$ where $\ell_{j,\varphi_{i-1}(\ell)}$ is a multiple of $\delta_{j,\varphi_{i-1}(\ell)}$, and therefore of $\delta_{j,\ell}$ since $\varphi_{i-1}(\ell) \geq \ell$. Moreover $\delta_{j,n} x_j \in \mathbb{Z}$ so that

$$\frac{\delta_{p,n}}{\delta_{p,\ell}} \ell_{j,\varphi_{i-1}(\ell)} x_j = \left(\frac{\delta_{p,n}/\delta_{j,n}}{\delta_{p,\ell}/\delta_{j,\ell}} \right) \left(\frac{\ell_{j,\varphi_{i-1}(\ell)}}{\delta_{j,\ell}} \right) (\delta_{j,n} x_j) \in \mathbb{Z}$$

since $\ell \leq n$, by assumption on the divisors $\delta_{t,n}$. Therefore we have $L_{\varphi_{i-1}(\ell)}(P) \in \frac{\delta_{p,\ell}}{\delta_{p,n}} \mathbb{Z}$, and the upper bound (3.11) implies that this rational number is zero for any $i \in \{1, \ldots, k\}$. Using Eq. (3.10) this yields the following upper bound on $|L_{\varphi_{i-1}(\ell)}(M)|$ (where we let $M = \sum_{j=1}^{k} \lambda_j e_j$):

$$|L_{\varphi_{i-1}(\ell)}(M)| = |L_{\varphi_{i-1}(\ell)}(u)| \leq \left(\frac{Q_\ell}{Q_n} \right)^{d_p} Q^{-\varepsilon/2}.$$

Combining this upper bound with Eq. (3.8) yields, for any $j \in \{1, \ldots, k\}$:

$$\begin{aligned}
|\lambda_j L_{\ell-1}(e_j)| &\leq \left(1 + \frac{1}{k} + \frac{1}{k^2}\right) \sum_{i=1}^{k} \left(\frac{Q_\ell}{Q_n} \right)^{d_p} Q^{-\varepsilon/2} Q_{\varphi_{i-1}(\ell)}^{\tau_j + o(1)} Q_{\ell-1}^{-\tau_j + o(1)} \\
&\leq Q_\ell^{d_p + \tau_j((1+\varepsilon_1)^{i-1} - 1) + o(1)} Q_n^{-d_p} Q^{-\varepsilon/2} \text{ using Eq. (3.6)} \\
&\leq Q_\ell^{d_p + \varepsilon/4 + o(1)} Q_n^{-d_p} Q^{-\varepsilon/2} \text{ using the assumption } \tau_j \leq \tau_1 \text{ and Eq. (3.5)} \\
&\leq \left(\frac{Q_\ell}{Q_n} \right)^{d_p} Q^{-\varepsilon/4 + o(1)} \leq \frac{\delta_{p,\ell}}{3k\delta_{p,n}}
\end{aligned}$$

since $Q_\ell \leq Q_n = Q^{1+o(1)}$ and $\frac{\delta_{p,\ell}}{\delta_{p,n}} = \frac{Q_\ell^{d_p + o(1)}}{Q_n^{d_p + o(1)}}$. This contradicts the minimality of ℓ in Eq. (3.9), thereby concluding the proof of (iii).

4 Consequences and related results

In this section we state and prove consequences of our main result (§§4.1 and 4.2), and mention Diophantine applications (§4.4). We also prove in §4.3 an analogous result, in the spirit of Siegel's linear independence criterion.

Throughout this section we restrict to the setting of Theorem 2, omitting for simplicity the refinements of Theorem 3 (eventhough they could have been adapted here).

Author's personal copy

4.1 Distance to integers

In this section we state corollaries of our criterion dealing with linear forms which are close to integers (rather than close to 0), as in Khintchine–Groshev's theorem for instance. In particular we deduce from Theorem 3 a result (namely Corollary 3 below) analogous to Nesterenko's linear independence criterion but which applies to sequences of simultaneous approximations of real numbers with the same denominator. This result is related to type II Padé approximation problems, in the same way as Nesterenko's criterion is related to type I problems. In this respect, Theorem 3 makes a bridge between the latter and the former: it is related to Padé approximation problems intermediate between type I and type II (see for instance [24]).

To begin with, let us state Theorem 2 in a *dual* way, namely in terms of $C_1, \dots, C_p \in \mathbb{R}^k$ rather than $e_1, \dots, e_k \in \mathbb{R}^p$.

Theorem 4 *Let $C_1, \dots, C_p \in \mathbb{R}^k$, with $k, p \geq 1$.*

Let $\tau_1, \dots, \tau_k$ and $(Q_n)_{n\geq 1}$ be as in Theorem 2.

For any $n \geq 1$, let $\ell_{1,n}, \dots, \ell_{p,n} \in \mathbb{Z}$ be such that, as $n \to \infty$:

$$\max_{1\leq i\leq p} |\ell_{i,n}| \leq Q_n^{1+o(1)} \text{ and } \ell_{1,n}C_1 + \cdots + \ell_{p,n}C_p = \begin{pmatrix} \pm Q_n^{-\tau_1+o(1)} \\ \vdots \\ \pm Q_n^{-\tau_k+o(1)} \end{pmatrix} \tag{4.1}$$

where the ± signs can be independent from one another. Then:

(i) *The rank of the family of vectors $C_1, \dots, C_p$ in $\mathbb{R}^k$, considered as a $\mathbb{Q}$-vector space, is greater than or equal to $k + \tau_1 + \cdots + \tau_k$.*

(ii) *For any non-zero linear form $\chi : \mathbb{R}^k \to \mathbb{R}$ there exists $i \in \{1, \dots, p\}$ such that $\chi(C_i) \notin \mathbb{Q}$.*

(iii) *Let $\varepsilon > 0$, and Q be sufficiently large in terms of ε. Let $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, not all zero, be such that $|\lambda_j| \leq Q^{\tau_j - \varepsilon}$ for any $j \in \{1, \dots, k\}$. Then denoting by χ the linear map $\mathbb{R}^k \to \mathbb{R}$ defined by $\chi(x_1, \dots, x_k) = \lambda_1 x_1 + \cdots + \lambda_k x_k$, we have*

$$\mathrm{dist}\Big((\chi(C_1), \dots, \chi(C_p)), \mathbb{Z}^p \setminus \{(0, \dots, 0)\}\Big) \geq Q^{-1-\varepsilon}$$

where $\mathrm{dist}(y, \mathbb{Z}^p \setminus \{(0, \dots, 0)\})$ is the minimal distance of $y \in \mathbb{R}^p$ to a non-zero integer point.

This result is just a translation of Theorem 2. Indeed let us consider the matrix $M \in \mathrm{Mat}_{k,p}(\mathbb{R})$ of which $C_1, \dots, C_p$ are the columns. We denote by $e_1, \dots, e_k \in \mathbb{R}^p$ the rows of M. Then assumption (4.1) means that the linear form $L_n = \ell_{1,n}X_1 + \cdots + \ell_{p,n}X_p$ on $\mathbb{R}^p$ is small at the points $e_1, \dots, e_k$. It is not difficult to see that (ii) and (iii) in Theorem 4 are respectively equivalent to (ii) and (iii) in Theorem 2, because $(\chi(C_1), \dots, \chi(C_p)) = \lambda_1 e_1 + \cdots + \lambda_k e_k$. We remark also that assuming $k \leq p - 1$ in Theorem 2 is not necessary; it has not been used in the proof. This upper bound follows from (ii), so that it is actually a consequence of the other assumptions.

Author's personal copy

Let us focus now on an important special case of Theorem 4, related to Padé approximation: when C_1, …, C_k is the canonical basis of $\mathbb{R}^k$. This happens in all practical situations mentioned in §4.4 below: indeed Padé approximation provides linear combinations of $C_{k+1}, \dots, C_p$ which are very close to $\mathbb{Z}^k$. In this case, in (ii) the interesting point is when the linear form $\chi(x_1, \dots, x_k) = \lambda_1 x_1 + \cdots + \lambda_k x_k$ has rational coefficients λ_j; then we have $\chi(C_i) \notin \mathbb{Q}$ for some $i \in \{k+1, \dots, p\}$. An analogous remark holds for (iii); both are more easily stated as follows, in terms of e_1,…, e_k. We denote by $\|\cdot\|$ any fixed norm on $\mathbb{R}^{p-k}$.

Corollary 2 *Under the assumptions of Theorem 2, suppose that for any $j \in \{1, \dots, k\}$ we have $e_j = (0, \dots, 0, 1, 0, \dots, 0, e'_j)$ with $e'_j \in \mathbb{R}^{p-k}$, where the 1 is in jth position.*

Then no non-trivial $\mathbb{Q}$-linear combination of $e'_1, \dots, e'_k$ belongs to $\mathbb{Q}^{p-k}$. In addition, let $\varepsilon > 0$, and Q be sufficiently large in terms of ε. Let $\lambda_1, \dots, \lambda_k \in \mathbb{Z}$, not all zero, be such that $|\lambda_j| \le Q^{\tau_j - \varepsilon}$ for any $j \in \{1, \dots, k\}$. Then for any $S \in \mathbb{Z}^{p-k}$ we have

$$\|\lambda_1 e'_1 + \cdots + \lambda_k e'_k - S\| \ge Q^{-1-\varepsilon}.$$

This corollary is a measure of linear independence of the vectors $e'_1, \dots, e'_k$ and those of the canonical basis of $\mathbb{Z}^{p-k}$. It can be weakened by assuming $|\lambda_j| \le Q^{\tau-\varepsilon}$ for any $j \in \{1, \dots, k\}$, where $\tau = \min(\tau_1, \dots, \tau_k)$ (as in Theorem 5 below). Then a *measure of non-discreteness* (in the sense of [15]) is obtained for the lattice $\mathbb{Z}e'_1 + \cdots + \mathbb{Z}e'_k + \mathbb{Z}^{p-k}$, which has rank p. In the examples (4.2), (4.3) and (4.4) considered in §4.4 below, the matrix with columns C_{k+1}, …, C_p is symmetric (with $p = 2k$), so that this lattice is exactly $\mathbb{Z}C_1 + \cdots + \mathbb{Z}C_p$ (using the fact that C_1, …, C_k is the canonical basis of $\mathbb{R}^k$).

This case $k = p/2$ lies "in the middle" between $k = 1$, which corresponds to type I Padé approximation and Nesterenko's original criterion, and $k = p - 1$, which corresponds to type II Padé approximation. In the latter case, Corollary 2 yields the following result by letting $\xi_j = -e'_j$.

Corollary 3 *Let $k \ge 1$, and $\xi_1, \dots, \xi_k \in \mathbb{R}$.*

Let $\tau_1, \dots, \tau_k > 0$ be pairwise distinct real numbers.

Let $(Q_n)_{n \ge 1}$ be an increasing sequence of positive integers, such that $Q_{n+1} = Q_n^{1+o(1)}$.

For any $n \ge 1$, let $\ell_{1,n}, \dots, \ell_{k,n}, \ell_{k+1,n} \in \mathbb{Z}$ be such that

$$\max_{1 \le i \le k+1} |\ell_{i,n}| \le Q_n^{1+o(1)}$$

and

$$|\ell_{k+1,n}\xi_j - \ell_{j,n}| = Q_n^{-\tau_j + o(1)} \text{ for any } j \in \{1, \dots, k\}.$$

Then:

(i) *The numbers 1, $\xi_1, \dots, \xi_k$ are $\mathbb{Q}$-linearly independent.*

Author's personal copy

(ii) *Let $\varepsilon > 0$, and Q be sufficiently large (in terms of ε). Then for any $(a_0, a_1, \dots, a_k) \in \mathbb{Z}^{k+1} \setminus \{(0, \dots, 0)\}$ with $|a_j| \le Q^{\tau_j - \varepsilon}$ for any $j \in \{1, \dots, k\}$, we have:*

$$|a_0 + a_1\xi_1 + \cdots + a_k\xi_k| \ge Q^{-1-\varepsilon}.$$

We have not found this statement in the literature; see however [8] (p. 98), [16] (Lemma 2.1) or [17] (Lemma 6.1) for related results, which are probably closer to Siegel's criterion than to Nesterenko's (see §4.3 below).

4.2 Upper bound on a Diophantine exponent

Given a subspace F of $\mathbb{R}^p$, and a non-zero point $P \in \mathbb{R}^p$, we denote by $\mathrm{Dist}(P, F)$ the projective distance of P to F, seen in $\mathbb{P}_{p-1}(\mathbb{R})$. Several definitions may be given, all of them equivalent up to multiplicative constants (see for instance [23]); we choose $\mathrm{Dist}(P, F) = \frac{\|u\|}{\|P\|}$ where u is the orthogonal projection of P on $F^\perp$ (that is, P can be written as $u + f$ with $u \in F^\perp$ and $f \in F$), and $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^p$.

The following result is a consequence of Theorem 2.

Theorem 5 *Under the assumptions of Theorem 2, let $\tau = \min(\tau_1, \dots, \tau_k)$ and $F = \mathrm{Span}_{\mathbb{R}}(e_1, \dots, e_k)$. Then for any $\varepsilon > 0$ and any $P \in \mathbb{Z}^p \setminus \{(0, \dots, 0)\}$ we have:*

$$\mathrm{Dist}(P, F) \ge \|P\|^{-1-\frac{1}{\tau}-\varepsilon}$$

provided $\|P\|$ is sufficiently large in terms of ε.

It is important to notice that Theorem 5 is *not* optimal, since it involves only $\min(\tau_1, \dots, \tau_k)$. It is specially interesting when $\tau_1, \dots, \tau_k$ are close to one another.

The interest of Theorem 5 is that it can be written as an upper bound on a Diophantine exponent which measures the approximation of F by points of $\mathbb{Z}^p$ (see [4,19,23]).

Proof of Theorem 5 Using assertion (ii) of Theorem 2, we see that $(e_1, \dots, e_k)$ is a basis of F. Since F is finite-dimensional, all norms on F are equivalent: there exists $\kappa > 0$ such that, for any $f = \lambda_1 e_1 + \cdots + \lambda_k e_k \in F$ (with $\lambda_j \in \mathbb{R}$), we have $\max|\lambda_j| \le \kappa\|f\|$.

Let $\varepsilon > 0$ be such that $\varepsilon < \tau$. Let Q_0 be such that assertion (iii) of Theorem 2 holds for any $Q \ge Q_0$; we assume that $\|P\| \ge Q_0^{\tau-\varepsilon}/\kappa$. Letting $Q = (\kappa\|P\|)^{1/(\tau-\varepsilon)}$ we have $Q \ge Q_0$. Since $P \in \mathbb{Z}^p \setminus \{(0, \dots, 0)\}$, P does not belong to the set $\mathcal{C}(\varepsilon, Q)$ defined in assertion (iii). Now writing $P = \lambda_1 e_1 + \cdots + \lambda_k e_k + u$ with $\lambda_j \in \mathbb{R}$ and $u \in F^\perp$, we have

$$\max_{1 \le j \le k} |\lambda_j| \le \kappa\|\lambda_1 e_1 + \cdots + \lambda_k e_k\| \le \kappa\|P\| = Q^{\tau-\varepsilon}$$

so that $\|u\| > Q^{-1-\varepsilon}$. Using the definition of Q and that of $\mathrm{Dist}(P, F)$, this concludes the proof of Theorem 5.

Author's personal copy

4.3 Connection with a Siegel-type criterion

The following result is analogous to Theorem 2, but its proof is much easier. It relies on Siegel's ideas for linear independence (see for instance [8], p. 81–82 and 215–216, or [20], Proposition 4.1). Special cases of this proposition have already been used in Diophantine results (see §4.4 below).

Proposition 1 *Let $1 \le k \le p-1$, and $e_1, \dots, e_k \in \mathbb{R}^p$ be $\mathbb{R}$-linearly independent vectors.*

Let $(Q_n)_{n\ge 1}$ be an increasing sequence of positive integers, and for any $n \ge 1$, let $L_n^{(t)} = \ell_{1,n}^{(t)} X_1 + \cdots + \ell_{p,n}^{(t)} X_p$ be p linearly independent linear forms on $\mathbb{R}^p$ (for $1 \le t \le p$), with integer coefficients $\ell_{i,n}^{(t)}$ such that, as $n \to \infty$:

$$|L_n^{(t)}(e_j)| \le Q_n^{-\tau_j + o(1)} \text{ for any } j \in \{1, \dots, k\} \text{ and any } t \in \{1, \dots, p\},$$

where $\tau_1, \dots, \tau_k > 0$ are real numbers, and

$$\max_{\substack{1\le i\le p \\ 1\le t\le p}} |\ell_{i,n}^{(t)}| \le Q_n^{1+o(1)}.$$

Then:

(a) *Conclusions (i) and (ii) of Theorem 2 hold.*
(b) *Let $\varepsilon > 0$, and n be sufficiently large (in terms of ε). Let $\mathcal{C}_n$ denote the set of all vectors that can be written as $\lambda_1 e_1 + \cdots + \lambda_k e_k + u$ with:*

$$\begin{cases} \lambda_1, \dots, \lambda_k \in \mathbb{R} \text{ such that } |\lambda_j| \le Q_n^{\tau_j - \varepsilon} \text{ for any } j \in \{1, \dots, k\} \\ u \in (\operatorname{Span}_{\mathbb{R}}(e_1, \dots, e_k))^{\perp} \text{ such that } \|u\| \le Q_n^{-1-\varepsilon} \end{cases}$$

Then $\mathcal{C}_n \cap \mathbb{Z}^p = \{(0, \dots, 0)\}$.

The main difference with Theorem 2 is that we require here p linearly independent linear forms for any n (and we also assume $e_1, \dots, e_k$ to be $\mathbb{R}$-linearly independent). This makes the proof much easier, and enables one to get rid of several important assumptions of Theorem 2 (namely $Q_{n+1} = Q_n^{1+o(1)}$, $\tau_1, \dots, \tau_k$ pairwise distinct, and $|L_n(e_j)|$ not too small).

If $Q_{n+1} = Q_n^{1+o(1)}$ in Proposition 1 then in (b) we may replace Q_n with any Q, by letting n be such that $Q_n \le Q < Q_{n+1}$.

Proof of Proposition 1 To prove conclusion (i) of Theorem 2, let F be a subspace of $\mathbb{R}^p$ defined over $\mathbb{Q}$, of dimension d, which contains e_1, …, e_k. Let n be sufficiently large. Up to reordering $L_n^{(1)}$, …, $L_n^{(p)}$, we may assume the restrictions of $L_n^{(1)}$, …, $L_n^{(d)}$ to F to be linearly independent linear forms on F. Denoting by $(u_1, \dots, u_d)$ a basis of F consisting in vectors of $\mathbb{Z}^p$, the matrix $[L_n^{(t)}(u_j)]_{1\le t,j\le d}$ has a non-zero integer determinant. By making suitable linear combinations of the columns, the values $L_n^{(t)}(e_1)$, …, $L_n^{(t)}(e_k)$ appear and lead to the upper bound $Q_n^{d-k-\tau_1-\cdots-\tau_k+o(1)}$ on the absolute value of this determinant. This concludes the proof of (i) of Theorem 2.

Author's personal copy

To prove part (b) of Proposition 1 (which implies conclusion (ii) of Theorem 2), we let $P = \lambda_1 e_1 + \cdots + \lambda_k e_k + u \in \mathcal{C}_n \cap \mathbb{Z}^p$ be non-zero; then $L_n^{(t)}(P) \neq 0$ for some t, but $L_n^{(t)}(P) \in \mathbb{Z}$ and $|L_n^{(t)}(P)| < 1$. This concludes the proof of Proposition 1.

4.4 Diophantine applications

The main interest of Theorems 2 and 3 is that they provide (in conclusion (i)) a lower bound for the rank of $(C_1, \ldots, C_p)$. Such a lower bound (with k essentially equal to a^{ε}) implies Theorem 1, using a general lemma of linear algebra (see [10] for details). This kind of lower bounds (with $k \geq 2$) exists in the literature: for instance Gutnik has proved [14] that the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -2\log 2 \\ \zeta(2) \end{pmatrix}, \quad \begin{pmatrix} \zeta(2) \\ -3\zeta(3) \end{pmatrix} \tag{4.2}$$

are $\mathbb{Q}$-linearly independent in $\mathbb{R}^2$ (so that, for any $r \in \mathbb{Q}^\star$, at least one number among $\zeta(2) - 2r\log 2$ and $3\zeta(3) - r\zeta(2)$ is irrational). More recently he has obtained also [15] the $\mathbb{Q}$-linear independence of

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2\zeta(3) \\ 3\zeta(4) \end{pmatrix}, \quad \begin{pmatrix} 3\zeta(4) \\ 6\zeta(5) \end{pmatrix}. \tag{4.3}$$

In the same spirit, T. Hessami-Pilehrood has proved [18] that if q is greater than some explicit function of k then the following $2k$ vectors are $\mathbb{Q}$-linearly independent in $\mathbb{R}^k$:

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \tag{4.4}$$

$$\begin{pmatrix} \mathrm{Li}_1(\frac{-1}{q}) \\ \mathrm{Li}_2(\frac{-1}{q}) \\ \vdots \\ \mathrm{Li}_k(\frac{-1}{q}) \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} \binom{j-1}{j-1}\mathrm{Li}_j(\frac{-1}{q}) \\ \binom{j}{j-1}\mathrm{Li}_{j+1}(\frac{-1}{q}) \\ \vdots \\ \binom{j+k-2}{j-1}\mathrm{Li}_{j+k-1}(\frac{-1}{q}) \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} \binom{k-1}{k-1}\mathrm{Li}_k(\frac{-1}{q}) \\ \binom{k}{k-1}\mathrm{Li}_{k+1}(\frac{-1}{q}) \\ \vdots \\ \binom{2k-2}{k-1}\mathrm{Li}_{2k-1}(\frac{-1}{q}) \end{pmatrix}.$$

The same result holds with $1/q$ instead of $-1/q$; see also Gutnik's preprints cited in [18].

These results share two common features: they rely on a special case of Proposition 1, and they prove the linear independence of the full set of p vectors involved. Using Theorem 3 it should not be difficult to produce alternative proofs of these results,

Author's personal copy

in which only one sequence of small linear forms is constructed (instead of p linearly independent ones). This may lead to further generalizations: for instance no proof of Ball–Rivoal's lower bound (1.1) is known without using Nesterenko's criterion. Moreover, it should be possible also to obtain lower bounds for the rank of a family of vectors [like (4.2) or (4.3) up to $\zeta(a)$, or (4.4) with smaller values of q] eventhough the present methods fail to prove the linear independence of the full set.

At last we would like to mention that during the submission process of the present paper, several results of the same flavour have been obtained by Dauguet [6] and applied to zeta values by Dauguet and Zudilin [7].

Acknowledgments The author has been partially supported by Agence Nationale de la Recherche (project HAMOT, ref. ANR 2010 BLAN-0115-01).

References

1. Apéry, R.: Irrationalité de $\zeta(2)$ et $\zeta(3)$, in Journées Arithmétiques (Luminy, 1978), Astérisque, no. 61, pp. 11–13 (1979)
2. Ball, K., Rivoal, T.: Irrationalité d'une infinité de valeurs de la fonction zêta aux entiers impairs. Invent. Math. **146**(1), 193–207 (2001)
3. Bourbaki, N.: Algèbre, ch. II, Hermann, third edn (1962)
4. Bugeaud, Y., Laurent, M.: On transfer inequalities in Diophantine approximation, II. Math. Z. **265**, 249–262 (2010)
5. Cassels, J.: An introduction to the geometry of numbers, Grundlehren der Math. Wiss., no. 99. Springer (1959)
6. Dauguet, S: Généralisations quantitatives du critère d'indépendance linéaire de Nesterenko, J. Théor. Nombres Bordeaux (to appear)
7. Dauguet, S., Zudilin, W.: On simultaneous diophantine approximations to $\zeta(2)$ and $\zeta(3)$. J. Number Theory **145**, 362–387 (2014)
8. Fel'dman, N., Nesterenko, Y.: Number theory IV, transcendental numbers, Encyclopaedia of Mathematical Sciences, no. 44. In: Parshin, A.N., Shafarevich, I.R. (eds.). Springer (1998)
9. Fischler, S.: Nesterenko's criterion when the small linear forms oscillate. Arch. der Math. **98**(2), 143–151 (2012)
10. Fischler, S.: Distribution of irrational zeta values, preprint, submitted. arxiv:1310.1685 (2013)
11. Fischler, S., Hussain, M., Kristensen, S., Levesley, J.: A converse to linear independence criteria, valid almost everywhere. Ramanujan J. (to appear)
12. Fischler, S., Rivoal, T.: Irrationality exponent and rational approximations with prescribed growth. Proc. Am. Math. Soc. **138**(8), 799–808 (2010)
13. Fischler, S., Zudilin, W.: A refinement of Nesterenko's linear independence criterion with applications to zeta values. Math. Ann. **347**, 739–763 (2010)
14. Gutnik, L.: On the irrationality of some quantities containing $\zeta(3)$, Acta Arith. **42**(3): 255–264, (1983) (in Russian); (translation in Amer. Math. Soc. Transl. **140**: 45–55 (1988))
15. Gutnik, L.: On linear forms with coefficients in $\mathbb{N}\zeta(1+\mathbb{N})$. In: Heath-Brown, D., Moroz, B. (eds.) Proceedings of the Session in analytic number theory and Diophantineequations (Bonn, 2002), Bonner Mathematische Schriften, no. 360, pp.1–45 (2003)
16. Hata, M.: Rational approximations to π and some other numbers. Acta Arith. **63**(4), 335–349 (1993)
17. Hata, M.: The irrationality of $\log(1+1/q)\log(1-1/q)$. Trans. Am. Math. Soc. **350**(6), 2311–2327 (1998)
18. Hessami Pilehrood, T.: Linear independence of vectors with polylogarithmic coordinates, Vestnik Moskov. Univ. Ser. I Mat. Mekh. [Moscow Univ. Math. Bull.] **54**(6): 54–56 [40-42] (1999)
19. Laurent, M.: On transfer inequalities in Diophantine approximation. In: Chen, W., Gowers, W., Halberstam, H., Schmidt, W., Vaughan, R. (eds.) Analytic Number Theory, Essays in Honour of Klaus Roth, pp. 306–314. Cambridge Univ. Press (2009)
20. Marcovecchio, R.: Linear independence of linear forms in polylogarithms. Annali Scuola Norm. Sup. Pisa V, no. 1, pp. 1–11 (2006)

Author's personal copy

21. Nesterenko, Y.: On the linear independence of numbers, Vestnik Moskov. Univ. Ser. I Mat. Mekh. [Moscow Univ. Math. Bull.] 40, no. 1, pp. 46–49 [69-74] (1985)
22. Rivoal, T.: La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs, C. R. Acad. Sci. Paris, Ser. I 331, no. 4, pp. 267–270 (2000)
23. Schmidt, W.: On heights of algebraic subspaces and Diophantine approximations. Ann. Math. **85**, 430–472 (1967)
24. Sorokin, V: A transcendence measure for π^2, Mat. Sbornik [Sb. Math.] **187**(12): 87–120 [1819-1852] (1996)