Astérisque

ALAIN PLAGNE

On the two-dimensional subset sum problem

Astérisque, tome 258 (1999), p. 375-409

<http://www.numdam.org/item?id=AST_1999__258__375_0>

Astérisque **258**, 1999, p. 375–409

ON THE TWO-DIMENSIONAL SUBSET SUM PROBLEM

by

Alain Plagne

Abstract. — We consider a system of two linear boolean equations. Using methods from analytic number theory, we obtain sufficient conditions ensuring the solvability of the system. This completes Freiman's work on the subject.

1. Introduction

In this paper, we are interested in considering the system of two linear equations

$$a_1x_1 + \cdots + a_mx_m = b, \tag{1}$$

where $a_i = (a_{i,1}, a_{i,2})$ and $b = (b_1, b_2)$ are in $\mathbb{Z}^2$ and the x_i's, the unknowns, restricted to be either 0 or 1: that is, we are only interested in the boolean system induced by (1). Our intention is to give sufficient conditions for the set of coefficients $A = \{a_1, \ldots, a_m\}$ and b to ensure the solvability of (1). Probabilistic considerations show that, if the a_i's are "well distributed" and if their number is large enough, we should have solutions for all b in the neighbourhood of $\sum_{i=1}^{m} a_i/2$ and, more precisely, that the distribution of the number of solutions must be Gaussian: in fact, we are expecting a central limit theorem. So that here we investigate conditions ensuring a "good" distribution and then deduce the general case, that is, we describe the structure of A^*, the set of all sums $a_1x_1 + \cdots + a_mx_m$ with boolean unknowns.

The corresponding one-dimensional problem has been much studied in the past recent years from this point of view (see for example **[F80, AF88, EF90, F93]** and **[C91b]** for a complete bibliography). It has been shown that A^* is a collection of arithmetical progressions with the same difference. Each of these papers uses methods coming from analytic number theory, in the vein introduced in the 80's by Freiman (in the first quoted paper), essentially the principle of the circle method.

1991 ***Mathematics Subject Classification.*** — 11P99, 11P55, 11H06.

Key words and phrases. — additive number theory, structure theorem, subset sum, two-dimensional subset sum problem, geometry of numbers, integer points, Farey dissection, convex set.

Freiman began to generalize these results in two dimensions **[F96]** but some details remained obscure (computations on page 143 for example). A little later, Chaimovich **[C91a]** tried to generalize this in higher dimensions but some algorithmical problems arose in these cases (see, for example, our counterexample in section 2.3 to the extension of Proposition 4 stated in **[C91a]**). Our goal here is to make clear the situation. We complete, correct and improve in some places Freiman's **[F96]**. In addition, the results given here are in an explicit form, because of the opportunity they offer to design algorithms. However the constants for which we prove the theorems are still far from being the best one could expect.

For the sake of completeness, the present paper is self-contained except for very classical tools (as, for example, Farey dissection) for which we refer as usual to **[HW]**.

In this paper we shall use the following notation: if u is in $\mathbb{R}^2$, we denote by u_1 and u_2 its coordinates with respect to the canonical basis (ϵ_1, ϵ_2) and by O the origin point. The e function is, as usual, defined by $e(t) = \exp(2\pi i t)$. For a real t, $||t||$ will denote the distance between t and $\mathbb{Z}$ and $[t]$ its integer part. The usual Euclidean scalar product is denoted simply with a point and the Lebesgue measure is denoted by μ. Finally, the volume of a fundamental parallelogram of any lattice Γ is denoted Vol Γ.

When $k, l \geq 1$ (in order to deal with really two-dimensional problems), we denote $P_{k,l}$ the integer rectangle

$$P_{k,l} = ([-k,k] \times [-l,l]) \cap \mathbb{Z}^2$$

and v its "volume", $v = (2k+1)(2l+1)$, that is, the number of integer points of $P_{k,l}$. In the sequel, A will denote a set of $m = |A|$ different integer points, $A = \{a_1, \dots, a_m\}$ and $J(b)$ the number of solutions of (1). We write $M = \sum_{i=1}^m a_i/2$ and

$$V = \begin{pmatrix} V_1^2 & V_{12} \\ V_{12} & V_2^2 \end{pmatrix},$$

where we have put $V_{12} = \sum_{j=1}^m a_{j,1}a_{j,2}$ and $V_i^2 = \sum_{j=1}^m a_{j,i}^2$ for $i = 1, 2$.

We denote by q_V the quadratic form naturally associated to this matrix $q_V(x) = \sum_{j=1}^m (a_j.x)^2$ $(x \in \mathbb{R}^2)$, and by $q_{V^{-1}}$ that one associated to V^{-1} that is $q_{V^{-1}}(x) = \frac{1}{\det V} \sum_{j=1}^m \det^2(a_j, x)$. Finally, we define the constants

$$k_1 = 25, \qquad k_2 = 6, \qquad k_4 = 189912,$$
$$k_5 = 100k_1 = 2500, \quad k_6 = 100k_2 = 600,$$
$$k_8 = \max(10k_6, k_5) = 6000, \qquad k_9 = \frac{9}{20},$$

and $k_3 = k_7$ being any constant $< 1/2$.

Our aim is to prove the following three Theorems:

Theorem 1. — *Let $A \subset P_{l_1,l_2}$ and $v = (2l_1+1)(2l_2+1)$. Assume*

$$|A| \geq k_1 v^{2/3} \log^{1/3} v \tag{2}$$

and that for each integer lattice Γ different from $\mathbb{Z}^2$ we have

$$|A \setminus A \cap \Gamma| \geq k_2 v^{2/3} \log^{1/3} v, \tag{3}$$

then we have the following asymptotic equivalent (when $v \to +\infty$)

$$J(b) \sim \frac{2^{m+1}}{\pi\sqrt{\det V}} \exp\{-2q_{V^{-1}}(M-b)\}, \tag{4}$$

provided that $q_{V^{-1}}(M-b) \leq k_3 \log\log v - 4$.

Notice first that the density hypothesis (2) implies

$$\frac{v}{\log v} \geq k_1^3,$$

that implies

$$v \geq k_4. \tag{5}$$

The previous Theorem is slightly better than Freiman's Theorem 1 of **[F96]**, the main difference being that the size of domain of validity of (4) is increased by a factor $\log\log v$ tending to infinity with v. This result is the heart of this work, but this is not entirely satisfying because dealing only with rectangle cases. That is why it is generalized in the following form.

Theorem 2. — *Let C be a compact convex set in $\mathbb{R}^2$ containing O, E be its integer points, and A be a subset of E. Assume*

$$|A| \geq k_5|E|^{2/3}\log^{1/3}|E|$$

and that for each integer lattice Γ different from $\mathbb{Z}^2$, we have

$$|A \setminus A \cap \Gamma| \geq k_6|E|^{2/3}\log^{1/3}|E|, \tag{6}$$

then we have the following asymptotic equivalent (when $|E| \to +\infty$)

$$J(b) \sim \frac{2^{m+1}}{\pi\sqrt{\det V}} \exp\{-2q_{V^{-1}}(M-b)\},$$

provided that $q_{V^{-1}}(M-b) \leq k_7 \log\log|E| - 4$.

Once again, it is not completely satisfying because it deals only with "good" cases: those where the elements of A are "well distributed". The conclusion of this paper will be the following general result.

Theorem 3. — *Let C be a compact convex set in $\mathbb{R}^2$ containing O, E be its integer points, and A be a subset of E. Assume $|A| \geq k_8|E|^{2/3}\log^{1/3}|E|$ and that for each line D such that $O \in D$, one has*

$$|A \cap D| < k_9|A|. \tag{7}$$

Then there exists a lattice Λ_0 such that, if A' stands for $A \setminus A \cap \Lambda_0$, one has $|A'| \leq |A \cap \Lambda_0|$ and

$$A'^* + (\Lambda_0 \cap F) \subset A^*,$$

where $F = \{x \in \mathbb{Z}^2, q_{W^{-1}}(M'-x) \leq k_7 \log\log(|A|/2) - 4\}$ and $W(x) = \sum_{a\in A'}(a.x)^2$, $M' = \sum_{a\in A'} a/2$.

This is a structural theorem because it describes how the set A^* is made, at least locally. It is a powerful result in order to design algorithms, as it has already been done in the one-dimensional case (see for example **[CFG89]**).

We notice that hypothesis (7) is in fact not very restrictive: it ensures that our set A is an essentially two-dimensional set. If that condition is not fulfilled, we have the possibility to treat our problem as a one-dimensional one, by forgetting some points and this is even much simpler.

Acknowledgments. — Many thanks go to G.A. Freiman for his kind help during the preparation of this paper. I would like also to thank J.-M. Deshouillers for his advice.

2. Preliminary lemmas

We begin this section by quoting some inequalities (whose validity can easily be seen by using, for instance, some Taylor-Lagrange's inequalities). For any real t, if $0 \leq |t| \leq 1/2$, we have

$$|1 + e(t)| \leq 2\exp(-\pi^2 t^2/2), \tag{8}$$

and if $|t| \leq \pi/2$,

$$0 \leq 1 - \exp(t^2/2)\cos t \leq (2t/\pi)^4. \tag{9}$$

Finally, for reals $(\epsilon_i)_{1 \leq i \leq n}$ between 0 and 1, we have, with a trivial induction argument,

$$\prod_{i=1}^{n}(1 - \epsilon_i) \geq 1 - \sum_{i=1}^{n} \epsilon_i. \tag{10}$$

Now we present several propositions that we shall need in the sequel.

2.1. Arithmetical lemmas.

— Here, we give two results concerning the number of solutions of a Diophantine inequality.

Lemma 1. — *Let a, b, ϵ be real numbers and k, n be integers such that $0 < |a|k < 1$ and $\epsilon < (1 - k|a|)/2$. Then we have*

$$|\{x \in \mathbb{N}, n \leq x \leq n + k : ||ax + b|| \leq \epsilon\}| \leq 1 + [2\epsilon/|a|].$$

Proof. — Without loss of generality we may assume $a > 0$ and write $u_s = as + b$. This is a strictly increasing sequence. Let s_1 be the smallest integer, with $n \leq s_1 \leq n + k$, such that $||u_{s_1}|| \leq \epsilon$ (if s_1 does not exist, then the cardinality studied is zero); we thus have $|u_{s_1} - e| \leq \epsilon$ for some integer e. Let s_2 be the largest integer satisfying $|u_{s_2} - e| \leq \epsilon$. We claim that $s_2 < t \leq n + k$ implies $||u_t|| > \epsilon$; indeed $|u_t - e| > \epsilon$ is clear by definition of s_2 and

$$u_t = u_{s_1} + (t - s_1)a \leq u_{s_1} + ka \leq e + \epsilon + ka < e + 1 - \epsilon.$$

Since $s_2 - s_1 = (u_{s_2} - u_{s_1})/a \leq 2\epsilon/a$, we get, for the cardinality studied, the desired upper bound. □

Now, we prove a result due to Freiman. We write here a complete proof, in view of the lack of details in Freiman's paper **[F96]**.

Proposition 1. — *Let n, k, P be integers, $0 \le P \le k$ and a, b be two reals. Assume $a = p/q + z$ with $(p,q) = 1$, $q \le P$, $1/2qk \le |z| \le 1/qP$. We have*

$$|\{x \in \mathbb{N}, n \le x \le n+k : ||ax+b|| \le P^{-1}\}| \le 3(4kP^{-1}+1).$$

Proof. — By just changing the value of b, the problem reduces to the case where $n = 0$. For $P \le 12$, the result is clear, so we assume from now on $P > 12$ and without loss of generality $z > 0$. The solutions of $||ax+b|| \le P^{-1}$ are clearly in bijection with those of the following problem, that we shall denote (P), consisting in finding $(x, x_0, t) \in \{0, \dots, k\} \times \{0, \dots, q-1\} \times \mathbb{Z}$ satisfying

$$\begin{cases} px \equiv x_0 \bmod q, \\ \left|\dfrac{x_0}{q} + zx + b - t\right| \le P^{-1}. \end{cases}$$

One can easily bound from above the cardinality, J, of the set of solutions of (P) as follows

$$\begin{aligned} J \le\ & |\{x_0 \mid \exists (x,t), (x,x_0,t) \text{ solution of (P)}\}| \\ & \times \max_{x_0} |\{t \mid \exists x, (x,x_0,t) \text{ solution of (P)}\}| \times \max_{x_0,t} |\{x \mid (x,x_0,t) \text{ solution of (P)}\}|. \end{aligned}$$

Now, write $|x_0/q + zx + b - t| \le P^{-1}$ in the following form

$$(11) \qquad -qP^{-1} - zxq - bq + tq \le x_0 \le qP^{-1} - zxq - bq + tq.$$

It implies, because $x \ge 0$, that x_0 belongs to $[-qP^{-1} - zxq - bq + tq, qP^{-1} - bq + tq]$. But t is an integer, $0 \le x \le k$ and x_0 stays in $\{0, \dots, q-1\}$, so x_0 belongs to $[-qP^{-1} - zkq - bq, qP^{-1} - bq] \bmod q$, which has length $2qP^{-1} + zkq$, this yields

$$|\{x_0 | \exists (x,t), (x,x_0,t) \text{ solution of (P)}\}| \le \inf([2qP^{-1} + zkq] + 1, q) \le \inf([zkq] + 3, q).$$

Now, the value of x_0 being given, equation (11) can be rewritten

$$-P^{-1} + x_0/q + zx + b \le t \le P^{-1} + x_0/q + zx + b,$$

and, as $0 \le x \le k$, one has

$$-P^{-1} + x_0/q + b \le t \le P^{-1} + x_0/q + zk + b,$$

thus t belongs to an interval of length $2P^{-1} + zk$, which implies

$$\max_{x_0} |\{t | \exists x, (x,x_0,t) \text{ solution of (P)}\}| \le [2P^{-1} + zk] + 1.$$

In the same vein, we can get

$$(12) \qquad \max_{x_0,t} |\{x | (x,x_0,t) \text{ solution of (P)}\}| \le [2/qPz] + 1.$$

Indeed, x_0 and t being given, we have

$$(-P^{-1} - x_0 q - b + t)/z \le x \le (P^{-1} - x_0 q - b + t)/z,$$

so the x's which are possible solutions are consecutive integers in an interval of length $2/Pz$. The condition $px \equiv x_0 \bmod q$ implies moreover that on a complete set of residues modulo q only one x can be solution. This proves (12).

We have finally, and in any case,

$$J \leq \inf([|z|kq] + 3, q)([2P^{-1} + |z|k] + 1)([2/qP|z|] + 1). \tag{13}$$

At this point, we have to distinguish two cases.

If $(1 + 2P^{-1})/2k \leq |z| \leq 1/qP$, equation (13) gives

$$J \leq q(2P^{-1} + 1 + |z|k)(1 + 2/|z|qP),$$

but, in view of the hypothesis, this is $\leq q(3|z|k)(3/|z|qP) = 9kP^{-1}$.

Now, if $|z| \leq (1 + 2P^{-1})/2k$, one has $2P^{-1} + |z|k < 2P^{-1} + (1 + 2P^{-1})/2 = 1/2 + 3P^{-1} < 1$, because $P > 12$, thus (13) implies

$$J \leq ([|z|kq] + 3)([2/qP|z|] + 1). \tag{14}$$

There are now three sub-cases according to the position of $|z|kq$.

If $|z|kq \geq 3/2$, (14) leads to

$$\begin{aligned} J &\leq (3 + |z|kq)(1 + 2/qP|z|) \\ &\leq (3|z|q)(3/qP|z|) = 9kP^{-1}, \end{aligned}$$

because one has, in every case $1 \leq 1/qP|z|$.

If now $1 \leq |z|kq \leq 3/2$, equation (14) yields

$$\begin{aligned} J &\leq ([3/2] + 3)(1 + 2/qP|z|) \\ &\leq 4(1 + 2kP^{-1}) \leq 12kP^{-1} + 3, \end{aligned}$$

because $kP^{-1} \geq 1$.

Finally, if $1/2 \leq |z|kq < 1$, in virtue of (14),

$$\begin{aligned} J &\leq ([|z|kq] + 3)(1 + 2/qP|z|) \\ &\leq 3(1 + 4kP^{-1}) \leq 12kP^{-1} + 3. \end{aligned}$$

This concludes the proof of Proposition 1. □

2.2. A two-dimensional "reverse-Cauchy-Schwarz" inequality.

— This section is devoted to the proof of an inequality used in **[F96]** without explanation. Our aim is to find a good lower bound for the ratio

$$\left(\sum_{j=1}^{m} (a_j.\alpha)^2\right)^2 \Big/ \sum_{j=1}^{m} (a_j.\alpha)^4 \tag{15}$$

which is naturally ≥ 1 and $\leq m$ (by the Cauchy-Schwarz inequality). We would like to "reverse" the Cauchy-Schwarz inequality, that is to find, for (15), a better lower bound than 1 (a power of m or $\log m$ for example). This is generally not possible, but here the a_j's have special properties which allow to get the desired result.

Let us consider the one-dimensional corresponding problem. Since α^4 can be factorized, the problem becomes to minimize

$$\left(\sum_{j=1}^{m} a_j^2\right)^2 \Big/ \sum_{j=1}^{m} a_j^4 \tag{16}$$

for distinct integers a_j's satisfying $1 \le a_j \le l$. It is easily seen that this ratio is

$$\ge \frac{\left(\sum_{j=1}^m a_j^2\right)^2}{l^2 \sum_{j=1}^m a_j^2} = \frac{\sum_{j=1}^m a_j^2}{l^2} \ge \frac{\sum_{j=1}^m j^2}{l^2} \sim \frac{m^3}{3l^2},$$

which is better than $O(1)$ as soon as $l^{2/3} = o(m)$.

This can be guessed in another way: if one tries to choose the a_j's such that (16) is near to 1, a natural idea (see below) is to take $a_j = j$ for $1 \le j \le m-1$ and $a_m = l$. This choice yields

$$\left(\sum_{j=1}^{m} a_j^2\right)^2 \Big/ \sum_{j=1}^{m} a_j^4 \asymp \frac{m^6 + l^4}{m^5 + l^4},$$

and this won't be $O(1)$ as soon as $l^4 = o(m^6)$, that is to say $l^{2/3} = o(m)$.

In dimension 2, the situation is not so clear but we will show that an analogous phenomenon happens. We begin by proving a preliminary lemma, which corresponds to a generalized one-dimensional case, for which we present two proofs: the first one will be direct while the second one corresponds to what we called the "natural idea" above. Although this second approach is much more intricate, we believe that the method could be efficient in some other contexts where the first one would not work.

The notation

$$\{a_1^{(e_1)}, \dots, a_n^{(e_n)}\}$$

is for the multi-set (that is the set "with repetition") composed with e_1 times a_1, e_2 times a_2, and so on.

Lemma 2. — *Let r, s be integers ≥ 1, $A \subset E = \{1^{(r)}, \dots, s^{(r)}\}$. Assume that*

$$|A| \ge c|E|^{2/3} \log^{1/3} |E| \tag{17}$$

for some constant c, then we have, for $k_{10} = 1/10$,

$$\left(\sum_{a\in A} a^2\right)^2 \ge k_{10} c^3 \log |E| \left(\sum_{a\in A} a^4\right). \tag{18}$$

First proof of Lemma 2. — We use the fact that

$$F(A) = \left(\sum_{a\in A} a^2\right)^2 \Big/ \left(\sum_{a\in A} a^4\right) \ge \left(\sum_{a\in A} a^2\right) \Big/ s^2. \tag{19}$$

Now, suppose first that $|A| \geq \sqrt{10}r$, then

$$F(A) \geq \frac{r(1^2 + 2^2 + \cdots + [|A|/r]^2)}{s^2} \geq \frac{r}{3s^2}\left[\frac{|A|}{r}\right]^3$$

which can be bounded from below by

$$\frac{r}{3s^2}\left(\frac{|A|}{r} - 1\right)^3 \geq \frac{1}{3}\left(1 - \frac{1}{\sqrt{10}}\right)^3 \frac{|A|^3}{r^2 s^2} \geq \frac{|A|^3}{10 r^2 s^2}.$$

Since $rs = |E|$, we get the lower bound $\frac{c^3}{10}\log|E|$.

Suppose now $|A| < \sqrt{10}r$, then

$$\sqrt{10}r > |A| \geq c(rs)^{2/3}\log^{1/3}|E|,$$

which furnishes

$$\frac{r}{s^2} > \frac{c^3}{10\sqrt{10}}\log|E|.$$

But (19) implies

$$F(A) \geq \frac{|A|}{s^2} \geq c\left(\frac{r}{s^2}\right)^{2/3}\log^{1/3}|E|$$

and thus

$$F(A) \geq \frac{c^3 \log|E|}{10}.$$

□

Now, we present the second method for obtaining a proof of Lemma 2 (in fact, the proof given here does not yield the same value for k_{10} but we did not try to optimize it). It begins with some definitions and a lemma.

Let E be a multi-set. If A is a sub-multi-set of E, a an element of A and b an element of $E \setminus A$, we denote by

$$A_a(b) = (A \setminus \{a\}) \cup \{b\},$$

the set obtained by replacing a by b in A.

Suppose E is a multi-set of reals and $\mathcal{F}$ a sub-family of the family of all sub-multi-sets of E. If A is a sub-multi-set of E belonging to $\mathcal{F}$ and a an element of A, we say that A is a-minimal relatively to $\mathcal{F}$ if for any b in $E \setminus A$ such that $b < a$, one has

$$A_a(b) \notin \mathcal{F}.$$

In the same way, we define A to be a-maximal relatively to $\mathcal{F}$ if for any b in $E \setminus A$ such that $b > a$, one has

$$A_a(b) \notin \mathcal{F},$$

and A is said to be a-extremal relatively to $\mathcal{F}$ if it is a-minimal or a-maximal relatively to $\mathcal{F}$. Finally, we say that A is $\mathcal{F}$-extremal if for any a in A, A is a-extremal. This can be restated in the following way: A is $\mathcal{F}$-extremal if for any $a \in A$ and $b, c \in E \setminus A$ such that $b < a < c$ then at least one of the sets $A_a(b), A_a(c)$ is not in $\mathcal{F}$. For example, if E is finite and $\mathcal{F} = \mathcal{P}(E)$, the $\mathcal{F}$-extremal sub-multi-sets are those in which the elements are accumulated on the extremities, with no "hole".

Lemma 3. — *Let t be any real, E be a finite multi-set of positive reals and $\mathcal{F}$ be any sub-family of the family of all sub-multi-sets of E. Assume the sub-multi-set B of E minimizes, on $\mathcal{F}$, the function D defined, for any $A \in \mathcal{F}$, by the formula*

$$D(A) = \Big(\sum_{a\in A} a^2\Big)^2 - t\Big(\sum_{a\in A} a^4\Big),$$

then B is $\mathcal{F}$-extremal.

Proof. — As E is finite, there is a minimum (on the sets belonging to $\mathcal{F}$) for D: so B always exists. Assume that B is not $\mathcal{F}$-extremal: it contains an element β such that there exists α, γ not in B and such that $0 \leq \alpha < \beta < \gamma$ holds. Denoting simply $B(\alpha)$ and $B(\gamma)$ the sets obtained by replacing β in B, respectively by α and γ, it follows, by hypothesis, that $B(\alpha)$ and $B(\gamma)$ belong to $\mathcal{F}$. As B is minimal for D on $\mathcal{F}$, one has

$$D(B(\alpha)) - D(B) \geq 0 \qquad \text{and} \qquad D(B(\gamma)) - D(B) \geq 0.$$

Denoting S the sum of squares of $B \setminus \{\beta\}$, this can be rewritten,

$$(S+\alpha^2)^2 - (S+\beta^2)^2 - t(\alpha^4 - \beta^4) \geq 0, \tag{20}$$

$$(S+\gamma^2)^2 - (S+\beta^2)^2 - t(\gamma^4 - \beta^4) \geq 0. \tag{21}$$

Introduce now the following notations:

$$X = \alpha^4, \qquad Y = \beta^4, \qquad Z = \gamma^4,$$

and

$$F(u) = (S + u^{1/2})^2.$$

We have

$$F''(u) = -S/2u^{3/2},$$

which is strictly negative: therefore F is strictly concave. But equations (20) and (21) show

$$\frac{F(Y)-F(X)}{Y-X} \leq t \leq \frac{F(Z)-F(Y)}{Z-Y},$$

and that is not possible for a strictly concave function in view of $X < Y < Z$. □

Second proof of Lemma 2. — Let $\mathcal{F}$ be the set of all sub-multi-sets of E satisfying (17). Assume that (18) is proven for every $\mathcal{F}$-extremal set, then by Lemma 3, (18) is proven for every sub-multi-set of E belonging to $\mathcal{F}$ and we are done. Thus we only have to check that (18) holds for $\mathcal{F}$-extremal sets. That is what we do now, after having noticed that conditions (17) and $|E| \geq |A|$ imply

$$|A| \geq c^3 \log |E|.$$

Thus it is enough to get a lower bound with $k_{10}|A| \left(\sum_{a\in A} a^4\right)$ in the right-hand side of equation (18).

As above, we define the ratio

$$F(A) = \Big(\sum_{a\in A} a^2\Big)^2 \Big/ \sum_{a\in A} a^4.$$

We have to investigate the cases where A is of the following form

$$A = \{1^{(r)}, \ldots, (a-1)^{(r)}, a^{(x_a)}, (s-b)^{(x_b)}, (s-b+1)^{(r)}, \ldots, s^{(r)}\},$$

with $a \geq 1$, $0 \leq x_a, x_b \leq r$, b being possibly zero. We have (using elementary tools)

$$\begin{aligned} F(A) &= \frac{(r(1^2+\cdots+(a-1)^2)+x_a a^2+x_b(s-b)^2+r(s^2+\cdots+(s-b+1)^2))^2}{(r(1^4+\cdots+(a-1)^4)+x_a a^4+x_b(s-b)^4+r(s^4+\cdots+(s-b+1)^4))} \\ (22) \qquad &\geq \frac{(r(a-1)^3/3+x_a a^2+x_b(s-b)^2+rbs^2/3)^2}{r(a-1)^5+x_a a^4+x_b(s-b)^4+rbs^4}. \end{aligned}$$

Furthermore, the cardinality $|A|$ verifies:

$$(23) \qquad |A| = r(a+b-1)+x_a+x_b.$$

Consider now the following sub-cases.

(1) If $a=1$, equation (22) shows that

$$(24) \qquad F(A) \geq \frac{(x_a+x_b(s-b)^2+rbs^2/3)^2}{x_a+x_b(s-b)^4+rbs^4}.$$

(1a) If $b=0$ then equation (24) produces

$$F(A) \geq \frac{(x_a+x_b s^2)^2}{x_a+x_b s^4}.$$

(1a1) If $x_a \geq x_b s^4$, we get

$$2F(A) \geq x_a + \frac{x_b^2 s^4}{x_a} \geq x_a + \frac{x_b^2}{x_a} \geq \sup(x_a, x_b) \geq \frac{|A|}{2}$$

because of (23). Thus $F(A) \geq |A|/4$.

(1a2) If $x_a \leq x_b s^4$, we get

$$F(A) \geq \frac{x_a^2+x_b^2 s^4}{2x_b s^4} \geq \begin{cases} x_b/2 \\ \sqrt{\dfrac{x_a^2 x_b^2 s^4}{x_b^2 s^8}} = \dfrac{x_a}{s^2}, \end{cases}$$

the second lower bound following from the arithmetico-geometric inequality.

If $x_b \geq |A|/2$, one has $F(A) \geq |A|/4$. Otherwise, as in (1a1), equation (23) implies $|A| = x_a + x_b \leq 2r$. Using (17), we have

$$2r \geq |A| \geq c|E|^{2/3}\log^{1/3}|E| \geq c(sr)^{2/3}\log^{1/3}|E|,$$

from which we deduce $8r \geq c^3 s^2 \log|E|$. Now, writing s^2 as $(s^2 s)^{2/3}$, we get

$$\begin{aligned} s^2 &\leq \left(\frac{8rs}{c^3\log|E|}\right)^{2/3} \\ &\leq \frac{4|E|^{2/3}}{c^2\log^{2/3}|E|}. \end{aligned}$$

Finally, we have (as $x_a \geq |A|/2$)

$$F(A) \geq \left(\frac{|A|}{2}\right)\left(\frac{c^2 \log^{2/3}|E|}{4|E|^{2/3}}\right) \geq \frac{c^3}{8}\log|E|.$$

(1b) Now, $b \geq 1$. Equation (24) gives

$$F(A) \geq \frac{(rbs^2/3)^2}{x_a + r(b+1)s^4} \geq \frac{(rbs^2/3)^2}{r(1+(b+1)s^4)} \geq \frac{b}{9(b+2)}rb \geq \frac{rb}{27},$$

in the same manner as above. Once again, using (23), we deduce $|A| \leq r(b+2) \leq 3rb$. Finally, in this case we have $F(A) \geq |A|/81$.

(2) If $a \geq 2$, we have

$$F(A) \geq \frac{(r(a-1)^3/3 + x_b(s-b)^2 + rbs^2/3)^2}{ra^5 + x_b(s-b)^4 + rbs^4}. \tag{25}$$

(2a) If $b = 0$, we get

$$F(A) \geq \frac{(r/3)^2(a-1)^6 + x_b^2 s^4}{ra^5 + x_b s^4}. \tag{26}$$

(2a1) If $ra^5 \geq x_b s^4$ then equation (26) implies

$$F(A) \geq \frac{(r/3)^2(a-1)^6}{2ra^5} = \frac{(a-1)r}{18}\left(\frac{a-1}{a}\right)^5 \geq \frac{(a-1)r}{576}.$$

But (23) implies $3(a-1)r \geq |A|$, so that finally $F(A) \geq |A|/1728$.

(2a2) If $ra^5 \leq x_b s^4$ then equation (26) yields

$$F(A) \geq \frac{(r(a-1)^3/3)^2 + x_b^2 s^4}{2x_b s^4}.$$

Applying once again the arithmetico-geometric inequality, we deduce $F(A) \geq r(a-1)^3/3s^2$. As $3r(a-1) \geq |A|$, using condition (17) we deduce the lower bound

$$(3(a-1)r)^3 \geq |A|^3 \geq (c(rs)^{2/3}\log^{1/3}|E|)^3 = c^3(rs)^2\log|E|,$$

thus

$$r(a-1)^3/s^2 \geq (c^3\log|E|)/27.$$

Finally, we have in this case $F(A) \geq (c^3\log|E|)/81$.

(2b) If $b \geq 1$, we have

$$F(A) \geq \frac{(r(a-1)^3/3 + rbs^2/3)^2}{ra^5 + r(b+1)s^4} = \frac{((a-1)^3/3 + bs^2/3)^2 r}{a^5 + (b+1)s^4}. \tag{27}$$

(2b1) If $a \geq b$, the cardinality equation (23) shows that $|A| \leq r(2a+1) \leq 5(a-1)r$.

(2b11) If $a^5 \geq (b+1)s^4$, we have successively

$$F(A) \geq \frac{r(a-1)^6/9}{2a^5} = \frac{(a-1)r}{18}\left(\frac{a-1}{a}\right)^5 \geq \frac{(a-1)r}{576} \geq \frac{|A|}{2880}.$$

(2b12) If $a^5 \leq (b+1)s^4$, we have

$$F(A) \geq r\frac{(a-1)^6/9 + b^2s^4/9}{2(b+1)s^4},$$

and after applying the arithmetico-geometric inequality

$$F(A) \geq \frac{b}{9(b+1)}\left(\frac{r(a-1)^3}{s^2}\right) \geq \frac{r(a-1)^3}{18s^2}.$$

Now, proceeding as in (2a2), we get that ($|A| \leq r(2a-1)+2r \leq r(2a+1) \leq 5r(a-1)$)

$$r(a-1)^3/s^2 \geq (c^3 \log|E|)/125,$$

and finally $F(A) \geq (c^3 \log|E|)/2250$.

(2b2) If $a \leq b$, using (27), we get

$$\begin{aligned} F(A) &\geq rb^2s^4/(9\{(b+1)s^4 + a^5\}) \\ &\geq rb^2/18(b+1) \geq rb/36. \end{aligned}$$

Once again (23) yields $3rb \geq |A|$, whence $F(A) \geq |A|/108$.

This completes this proof of Lemma 2. □

Before going a step further, it is interesting to notice that hypothesis (3) implies trivially the following:

(28) For each line D containing O, $|A \setminus A \cap D| \geq k_2 v^{2/3} \log^{1/3} v$,

since $D \cap \mathbb{Z}^2$ can be completed in some integral lattice different from $\mathbb{Z}^2$.

We are now able to deduce the following

Proposition 2. — *If $A \subset P_{l_1,l_2}$ satisfies $|A| \geq k_1 v^{2/3} \log^{1/3} v$ and hypothesis (3), then, for every $\alpha \in \mathbb{R}^2$, we have (recall $v = (2l_1+1)(2l_2+1)$)*

$$\left(\sum_{a\in A}(a.\alpha)^2\right)^2 \geq k_{11}\left(\sum_{a\in A}(a.\alpha)^4\right)\log v,$$

where

$$k_{11} = 1.12\ 10^{-3}.$$

as a consequence of

Proposition 3. — *If $A \subset P_{l_1,l_2}$ satisfies $|A| \geq k_1 v^{2/3} \log^{1/3} v$ and hypothesis (28), then, for every $\alpha \in \mathbb{R}^2$, we have*

$$\left(\sum_{a\in A}(a.\alpha)^2\right)^2 \geq k_{11}\left(\sum_{a\in A}(a.\alpha)^4\right)\log v.$$

Proof. — Let us first notice some facts. The formula is homogeneous and continuous with respect to α and symmetrical (as P_{l_1,l_2} is). Thus it suffices to prove it for every $\alpha = (p,q)$ with p,q positive integers sufficiently large and $\gcd(p,q) = 1$ (during this proof l_1 and l_2 are assumed to be fixed). Indeed the fractions q/p subject to these conditions are dense in $\mathbb{R}^+$.

In all this proof, we write $N = pl_1 + ql_2$ and assume, with no loss of generality, that

$$ql_2 \geq pl_1, \tag{29}$$

and

$$p \geq 2l_2 + 1, \tag{30}$$

$$q \geq 2l_1 + 1. \tag{31}$$

Then

$$\mathcal{S} = \{|a_j.\alpha|, a_j \in A\} \subset \{0^{(1)}, 1^{(2)}, \dots, N^{(2)}\} \subset \{0^{(2)}, 1^{(2)}, \dots, N^{(2)}\};$$

indeed $x.\alpha = t \in \mathbb{Z}$ is the equation of a line which can have at most one point in P_{l_1,l_2} because if

$$\begin{cases} px_1 + qx_2 = t, \\ py_1 + qy_2 = t, \end{cases}$$

then $q|y_1 - x_1$. This implies $x_1 = y_1$ as a consequence of $|x_1|, |y_1| \leq l_1 < q/2$ and then $x_2 = y_2$.

We now examine the value of $|n.\alpha|$ when $n \in P_{l_1,l_2}$. Take first $u, v \in \mathbb{Z}$ by Bezout Theorem such that $pu + qv = 1$. If $n.\alpha = t$, then there exists an integer e such that

$$n = t(u,v) + e(q,-p).$$

So, $n \in A$ implies that $|tu + eq| \leq l_1$ for some integer e, that is to say

$$\left\| t\frac{u}{q} \right\| \leq \frac{l_1}{q}. \tag{32}$$

We now distinguish two cases.

First case. — We assume that $(2l_2+1)^2 \geq 2l_1 + 1$ or that $q/p \leq 2v^{1/3}/3$. In the case where $(2l_2 + 1)^2 \geq 2l_1 + 1$, we get

$$(2l_1 + 1) \leq (2l_1 + 1)^{2/3}(2l_2 + 1)^{2/3} = v^{2/3},$$

and in the case where $q/p \leq 2v^{1/3}/3$, we get (using relation (29))

$$(2l_1 + 1) \leq (2l_1 + 1)^{1/2} \left(\frac{3q}{2p}(2l_2 + 1) \right)^{1/2} \leq v^{2/3}.$$

Here we have used $2l_1 + 1 \leq 3l_1 \leq 3ql_2/p \leq \frac{3q}{2p}(2l_2 + 1)$.

Let $k = [q/2] + 1$ and $P = [q/2l_1] < k$, we approximate u/q by an element α/β of the Farey dissection of order P:

$$\frac{u}{q} = \frac{\alpha}{\beta} + z,$$

with

$$(2k)^{-1} \leq q^{-1} \leq \beta|z| \leq P^{-1},$$

the lower bound being due to the fact that $u/q \neq \alpha/\beta$ because $\beta \leq P < q$ and $\gcd(u,q) = 1$. We can apply Proposition 1 that yields

$$\left| \left\{ -k \leq t \leq k : \left\| t\frac{u}{q} \right\| \leq P^{-1} \right\} \right| \leq 3(8kP^{-1} + 1) \leq 13(2l_1 + 1),$$

if q is large enough. Almost similarly (we have to consider separately the cases $t > 0$ and $t < 0$ but we get the upper bound $6(4kP^{-1}+1)$ and finally the same result), for any integer $0 \leq w \leq M = [N/k]$, we infer

$$\left|\left\{t \in \mathbb{Z}, wk \leq |t| \leq (w+1)k : \left\|t\frac{u}{q}\right\| \leq P^{-1}\right\}\right| \leq 13(2l_1+1);$$

thus, by putting $b_j = |a_j.\alpha|/k$ and

$$\mathcal{A}_w = \{j : w \leq b_j < w+1\},$$

for $w = 0, 1, \dots, M$, one has $|\mathcal{A}_w| \leq 13(2l_1+1)$. Now

$$\frac{\left(\sum_{j=1}^{m}(a_j.\alpha)^2\right)^2}{\left(\sum_{j=1}^{m}(a_j.\alpha)^4\right)} = \frac{\left(\sum_{j=1}^{m} b_j^2\right)^2}{\sum_{j=1}^{m} b_j^4} = \frac{\left(\sum_{w=0}^{M}\left(\sum_{j\in\mathcal{A}_w} b_j^2\right)\right)^2}{\sum_{j\in\mathcal{A}_0} b_j^4 + \sum_{w=1}^{M}\sum_{j\in\mathcal{A}_w} b_j^4}.$$

But, $\sum_{j\in\mathcal{A}_0} b_j^4 \leq |\mathcal{A}_0|$, thus

$$\sum_{w=1}^{M}\sum_{j\in\mathcal{A}_w} b_j^4 \geq \left|\bigcup_{w=1}^{M}\mathcal{A}_w\right| \geq |A|/2 \geq |\mathcal{A}_0| \geq \sum_{j\in\mathcal{A}_0} b_j^4,$$

because $|\mathcal{A}_0| \leq 13(2l_1+1) \leq 13v^{2/3} \leq |A|/2$ (this is due to the fact that $13 \leq (k_1/2)\log^{1/3} v$ for $v \geq k_4$). Therefore we obtain

$$\frac{\left(\sum_{j=1}^{m}(a_j.\alpha)^2\right)^2}{\left(\sum_{j=1}^{m}(a_j.\alpha)^4\right)} \geq \frac{\left(\sum_{w=1}^{M}|\mathcal{A}_w|w^2\right)^2}{2\sum_{w=1}^{M}(w+1)^4|\mathcal{A}_w|} \geq \frac{\left(\sum_{w=1}^{M}|\mathcal{A}_w|w^2\right)^2}{2^5\sum_{w=1}^{M}w^4|\mathcal{A}_w|} = \frac{F(C)}{2^5},$$

in the notations of the proof of Lemma 2 with $C \subset E = \{1^{(r)}, \dots, s^{(r)}\}$ and $r = 13(2l_1+1)$ and $s = M = [N/k]$. We have (q large), using inequality (29),

$$\begin{aligned} |E| = rs &\leq 13(2l_1+1)N/k \leq 52(2l_1+1)l_2 \leq 26v, \\ |E| &\geq 13(2l_1+1)(N/k-1) \geq 13(2l_1+1)(3l_2/2-1) \geq 2v. \end{aligned}$$

Thus $|E| \geq 2k_4$ that implies

$$\begin{aligned} |C| = |A| - |\mathcal{A}_0| \geq \frac{|A|}{2} \geq \frac{k_1}{2}v^{2/3}\log^{1/3} v &\geq \frac{k_1}{2}(|E|/26)^{2/3}\log^{1/3}(|E|/26) \\ &\geq \frac{k_1}{4(26)^{2/3}}|E|^{2/3}\log^{1/3}|E|. \end{aligned}$$

Consequently, thanks to Lemma 2, we get the lower bound

$$\frac{F(C)}{2^5} \geq \frac{k_{10}k_1^3}{1384448}\log|E| \geq \frac{k_{10}k_1^3}{1384448}\log v.$$

Second case. — We now consider degenerate cases, namely when $2l_1+1 \geq (2l_2+1)^2$ and $q/p \geq 2v^{1/3}/3$. This corresponds to cases where P_{l_1,l_2} is "thin" and α "almost orthogonal" to P_{l_1,l_2}. It requires a particular treatment.

We examine the case where

$$q/p \leq 2l_1 \tag{33}$$

and show that what has been done in the previous lines holds. We put $\epsilon = l_1/q$ and

$$k = [q(2l_2+1)^{2/3}/(2l_1+1)^{1/3}] \geq 1,$$

for large enough q. By using the Bezout relation, we see that

$$t\frac{u}{q} = t\frac{-v}{p} + \frac{t}{pq}.$$

We obtain

$$\begin{aligned}\left|\left\{-k \leq t \leq k : \left\|t\frac{u}{q}\right\| \leq \epsilon\right\}\right| &= \left|\left\{0 \leq t \leq 2k : \left\|\frac{tv}{p} + \frac{ku}{q} - \frac{t}{pq}\right\| \leq \epsilon\right\}\right| \\ &\leq \sum_{w=0}^{[2k/p]} \left|\left\{wp \leq t < (w+1)p : \left\|\frac{tv}{p} + \frac{ku}{q} - \frac{t}{pq}\right\| \leq \epsilon\right\}\right| \\ &\leq \sum_{w=0}^{[2k/p]} \left|\left\{wp \leq t < (w+1)p : \left\|\frac{tv}{p} + \frac{ku-w}{q}\right\| \leq \eta\right\}\right|\end{aligned}$$

where

$$\eta = \epsilon + q^{-1}.$$

But, as v is invertible modulo p, the number of solutions to $\left\|\frac{tv}{p} - c\right\| \leq \eta$ in a residue class modulo p is $\leq 2\eta p + 1$. Thus

$$\left|\left\{-k \leq t \leq k : \left\|t\frac{u}{q}\right\| \leq \epsilon\right\}\right| \leq (2\eta p + 1)(1 + 2k/p). \tag{34}$$

For q large enough, one has

$$\frac{k}{p} = \frac{\left[\frac{q(2l_2+1)^{2/3}}{(2l_1+1)^{1/3}}\right]}{p} \sim \frac{q(2l_2+1)^{2/3}}{p(2l_1+1)^{1/3}}$$

and this is

$$\geq (2v^{1/3}/3)\frac{(2l_2+1)^{2/3}}{(2l_1+1)^{1/3}} = 2(2l_2+1)/3 \geq 2.$$

Therefore for q large, $k/p \geq 1$. Concerning ηp, using the supplementary hypothesis (33), we have

$$2\eta p = 2\left(\frac{l_1+1}{q}\right)p \geq \frac{2l_1 p}{q} \geq 1,$$

thus (34) leads to

$$\left|\left\{-k \leq t \leq k : \left\|t\frac{u}{q}\right\| \leq \epsilon\right\}\right| \leq 4\eta p(3k/p) = 12\eta k$$
$$\leq 12\left(\frac{l_1+1}{q}\right)\left(\frac{(2l_2+1)^2}{2l_1+1}\right)^{1/3} q \leq 8v^{2/3}.$$

As above, we get the same result in the general case:

$$\left|\left\{wk \leq |t| < (w+1)k : \left\|t\frac{u}{q}\right\| \leq \epsilon\right\}\right| \leq 8v^{2/3},$$

and, with the same notation as before, we have

$$|E| \leq 8v^{2/3}(1+N/k) \leq 8v^{2/3}(1+2v^{1/3}) \leq 24v,$$

therefore we can conclude as previously.

Case $\alpha = (0,1)$. — To complete the result, we first establish it in the case where $\alpha = (0,1)$. In this case, we have

$$\frac{\left(\sum_{j=1}^{m}(a_j.\alpha)^2\right)^2}{\left(\sum_{j=1}^{m}(a_j.\alpha)^4\right)} = \frac{\left(\sum_{j\in J} a_{j,2}^2\right)^2}{\left(\sum_{j\in J} a_{j,2}^4\right)},$$

with $J = \{j : |a_{j,2}| \neq 0\}$. If $J_1 = \{|a_{j,2}|, j \in J\} \subset E = \{1^{(4l_1+2)}, \dots, l_2^{(4l_1+2)}\}$, hypothesis (28) applied to the line $\mathbb{R}e_1$ implies (since $v \geq |E|$)

$$|J_1| \geq k_2 v^{2/3} \log^{1/3} v \geq k_2 |E|^{2/3} \log^{1/3} |E|,$$

which permits us to apply Lemma 2 with $c = k_2$ and to conclude that the fraction is

$$\geq k_{10}k_2^3 \log|E| \geq \frac{k_{10}k_2^3}{2} \log v,$$

because $|E| = (4l_1+2)l_2 \geq 2v/3$.

Extension of the formula. — Now, we extend the formula by continuity in the neighbourhood of $\alpha = (0,1)$ to fill the gap, namely we have to show that for every $0 \leq \theta \leq 1/2l_1$, the relation holds for the vector $(\theta, 1)$. But for any $a_{j,1}$, $|a_{j,1}\theta| \leq 1/2$, thus if we denote by K the set of j's such that $a_{j,2} = 0$ and by J its complementary (on which $|a_{j,2}|/2 \leq |a_{j,1}\theta + a_{j,2}| \leq 3|a_{j,2}|/2$), we obtain

$$F = \frac{\left(\sum_{j=1}^{m}(a_{j,1}\theta + a_{j,2})^2\right)^2}{\sum_{j=1}^{m}(a_{j,1}\theta + a_{j,2})^4} \geq \frac{\left(\sum_{j\in K} a_{j,1}^2\theta^2 + \frac{1}{4}\sum_{j\in J} a_{j,2}^2\right)^2}{\sum_{j\in K} a_{j,1}^4\theta^4 + \left(\frac{3}{2}\right)^4 \sum_{j\in J} a_{j,2}^4}$$

$$\geq \frac{S_{K,2}^2\theta^4 + S_{J,2}^2}{81(S_{K,4}\theta^4 + S_{J,4})} = g(\theta^4)/81,$$

where $S_{K,2} = \sum_{j\in K} a_{j,1}^2, S_{J,2} = \sum_{j\in J} a_{j,2}^2, S_{K,4} = \sum_{j\in K} a_{j,1}^4, S_{J,4} = \sum_{j\in J} a_{j,2}^4$. This is a monotonic function g of θ^4. Thus

$$81F \geq \inf(g(0), g((1/2l_1)^4)).$$

We just estimated $g(0)$ (cf. Case $\alpha = (0,1)$), there remains to calculate $g((1/2l_1)^4)$,

$$g((1/2l_1)^4) \geq \frac{S_{K,2}'^2 + S_{J,2}'^2}{S_{K,4}' + S_{J,4}'},$$

where $S_{K,2}' = \sum_{j\in K}(a_{j,1}/2l_1)^2, S_{J,2}' = S_{J,2}, S_{K,4}' = \sum_{j\in K}(a_{j,1}/2l_1)^4, S_{J,4}' = S_{J,4}$. We have now to consider two different cases.

If $S_{K,4}' \geq S_{J,4}'$ (this implies $|K| \geq |J|$ and consequently $|K| \geq |A|/2$), then writing

$$g((1/2l_1)^4) \geq \frac{S_{K,2}'^2}{2S_{K,4}'^2} = \frac{S_{K,2}^2}{2S_{K,4}},$$

we can apply the result of Lemma 2, since the cardinality of K', the set of j's in K such that $a_{j,1} \neq 0$ verifies

$$|K'| \geq |K| - 1 \geq |A|/3 \geq \frac{k_1}{3}v^{2/3}\log^{1/3} v$$

and

$$\{|a_{j,1}|, j \in K'\} \subset \{1^{(2)}, \ldots, l_1^{(2)}\} \subset \{1^{(4l_2+2)}, \ldots, l_1^{(4l_2+2)}\} = E.$$

We have $2v/3 \leq (4l_2 + 2)l_1 = |E| \leq v$, thus we obtain

$$g((1/2l_1)^4) \geq \frac{k_{10}(k_1/3)^3}{2}\log(2v/3) \geq \frac{k_{10}k_1^3}{100}\log v.$$

If $S_{K,4}' \leq S_{J,4}'$ then

$$g((1/2l_1)^4) \geq \frac{S_{J,2}'^2}{2S_{J,4}'} = \frac{S_{J,2}^2}{2S_{J,4}}.$$

But $|J| \geq k_2 v^{2/3}\log^{1/3} v$ because of hypothesis (28) applied to the line $\mathbb{R}\epsilon_1$. As above $\{|a_{j,2}|, j \in J\} \subset E = \{1^{(4l_2+2)}, \ldots, l_1^{(4l_2+2)}\}$ and $2v/3 \leq |E| \leq v$ that allows to obtain the lower bound

$$g((1/2l_1)^4) \geq \frac{k_{10}k_2^3}{2}\log|E| \geq \frac{k_{10}k_2^3}{4}\log v.$$

All this computation show that, in fact, we can take any

$$k_{11} \leq \inf \left\{ \frac{k_1^3}{1384448}, \frac{k_2^3}{324} \right\} k_{10}$$

and ends the proof. □

2.3. Geometrical lemmas . — This section is devoted to the geometrical aspects of the problem. We complete Freiman's proof **[F96]** by studying every cases and improve some interesting intermediate results.

Through all this section we refer to **[C]** for extra information.

For C a compact convex body of $\mathbb{R}^2$, let us denote E its integer points, $E = C \cap \mathbb{Z}^2$. If Λ is a sub-lattice of $\mathbb{Z}^2$, we consider here $E \cap \Lambda$, which we assume to be two-dimensional, that is, not included in a line (this implies $|E| \geq 3$) and introduce some vocabulary and notation. If Δ is a line maximizing the cardinality $|\Delta \cap E \cap \Lambda|$, we write $\Delta \cap \mathbb{Z}^2 = \mathbb{Z}e_1$ for some e_1. Now, $\Delta \cap E \cap \Lambda = \{A_0 + k\alpha e_1, 0 \leq k \leq n\}$ for some point A_0, and α, n positive integers, because of the convexity of C. In the sequel, without loss of generality, we assume $A_0 = O$ and write $A = n\alpha e_1$. Next we choose e_2' completing e_1 in a $\mathbb{Z}^2$-basis, this is always possible. Now take β the unique (in view of $|\alpha\beta| =$ Vol Λ, the volume of a fundamental parallelogram of Λ) positive integer and γ' in $\mathbb{Z}$ such that

$$\Lambda = \alpha\mathbb{Z}e_1 + \mathbb{Z}(\beta e_2' + \gamma' e_1).$$

Define $u = \inf\{t \in \mathbb{Z}, \beta e_2' + te_1 \in E \cap \Lambda\}$ and $e_2 = e_2' + [u/\beta]e_1$. Then (e_1, e_2) is, as well, a $\mathbb{Z}^2$-basis and

$$\Lambda = \alpha\mathbb{Z}e_1 + \mathbb{Z}(\beta e_2 + \gamma e_1),$$

for some $\gamma \in \mathbb{Z}$. By definition of u, one can easily see that if $\beta e_2 + te_1 \in E \cap \Lambda$ then $t \geq \beta(u/\beta - [u/\beta]) \geq 0$. This remark will be needed in the sequel. Points of Λ are of the shape $(k\alpha + l\gamma)e_1 + l\beta e_2$ with $k, l \in \mathbb{Z}$. We note

$$d^+ = \max \{l | (k\alpha + l\gamma)e_1 + l\beta e_2 \in E \cap \Lambda \text{ for some } k\} \geq 0,$$

$$d^- = - \min \{l | (k\alpha + l\gamma)e_1 + l\beta e_2 \in E \cap \Lambda \text{ for some } k\} \geq 0.$$

Changing, if needed, e_2 in $-e_2$, one can assume that

$$d = \max\{d^+, d^-\} = d^+,$$

and since $E \cap \Lambda$ is not one-dimensional, $d \geq 1$. Clearly

$$|E \cap \Lambda| \leq (d^+ + d^- + 1)(n+1) \leq (2d+1)(n+1). \tag{35}$$

Finally, we note Δ_s the line $se_2 + \mathbb{R}e_1$ and define

$$c_s = |\Delta_s \cap E|,$$
$$c_s' = |\Delta_s \cap E \cap \Lambda|.$$

First we have to prove some preparatory lemmas.

Lemma 4. — *With the preceding conditions and notation, we have*

(i) $\Delta_{(n+1)\beta} \cap E \cap \Lambda$ *is empty,*

(ii) *If* $n \geq 2$*:* $d \leq n+3$.

(iii) *If* $n = 1$*:* $d \leq 3$.

Proof. — We first prove (i). Assume that $\Delta_{(n+1)\beta} \cap E \cap \Lambda$ is not empty. We can find

$$P = (k\alpha + (n+1)\gamma)e_1 + (n+1)\beta e_2,$$

a point in that intersection. Performing the Euclidean division of k by $n+1$, we find an integer r between 0 and n such that $k = (n+1)q + r$ for some integer q. Now, the convexity of C shows, on the one hand, that $r\alpha e_1$ belongs to E, because it is located between O and $n\alpha e_1$ and, on the other hand, that any integer point on the segment joining

$$P = (n+1)(\beta e_2 + (q\alpha + \gamma)e_1) + r\alpha e_1$$

and $r\alpha e_1$ is in E too; in particular, for each integer s belonging to $\{0, \dots, n+1\}$, the point $s(\beta e_2 + (q\alpha + \gamma)e_1)) + r\alpha e_1$ belongs to E. But now, these points are on a same line and their cardinality is $n+2$, which is impossible.

Let us now turn to (ii). Assume $n \geq 2$ and that there exists a positive integer j such that $\Delta_{(n+4+j)\beta} \cap E \cap \Lambda$ contains some point M. Convexity of C implies that the "full" triangle (AOM) is entirely in C. Let us denote L the length of the intersection of that triangle and $\Delta_{(n+1)\beta}$ (which is a segment). Application of Thales's Theorem gives

$$\frac{L}{n\alpha|e_1|} = \frac{3+j}{n+4+j},$$

that is

$$L = \frac{(3+j)n}{n+4+j}\alpha|e_1|,$$

an increasing expression with respect to n and j, which is therefore minimal when $n = 2$ and $j = 0$. It implies that $L \geq \alpha|e_1|$. But then $\Delta_{(n+1)\beta} \cap E \cap \Lambda$ contains at least one point and is subsequently not empty, contrarily to (i). Because each point of Λ is on some line $\Delta_{s\beta}$, one has $d \leq n+3$.

Now, let us see (iii): n is 1. If $d = 1$ or 2, there is nothing to prove. Assume we have $d \geq 3$ and choose M a point in $\Delta_{d\beta} \cap E \cap \Lambda$. Part (i) of the Lemma shows that $\Delta_{2\beta}$ does not intersect $E \cap \Lambda$, and then that the diameter of the intersection $\Delta_{2\beta} \cap C$ is less than $\alpha|e_1|$. Whence there exists a unique couple (S, T) of points of Λ verifying the properties

1. Non-void segment $\Delta_{2\beta} \cap C$ is included in the segment $[S, T]$,
2. $T = S + \alpha e_1$.

These points are of the following shape

$$S = (k\alpha + 2\gamma)e_1 + 2\beta e_2,$$
$$T = ((k+1)\alpha + 2\gamma)e_1 + 2\beta e_2 = S + \alpha e_1.$$

We now show that k is odd. In fact, if k were even, say $= 2l$, the point $S' = \frac{1}{2}S = (l\alpha + \gamma)e_1 + \beta e_2$ would be in Λ. Moreover, αe_1 and S' form a basis of Λ. But the "full" triangle (OAM) is contained in C, which forces the point M to be in the open strip (if not so, S or T would be in C) between the lines (OS) and (AT). Its coordinates are then of the form:

$$M = xS + y\alpha e_1,$$

with $x > 1$ and $0 < y < 1$. The point M can not belong to Λ, its coordinates not being entire (in the basis $(\alpha e_1, S')$). A contradiction.

The integer k is consequently odd, say $= 2l+1$. If $T' = \frac{1}{2}T$, as before, we get that T' and αe_1 generate Λ. Writing

$$M = xS + y\alpha e_1,$$

with $x > 1$ and $0 < y < 1$, we get $M = 2xT' + (y-x)\alpha e_1$. But, because M is in Λ, $2x$ and $y - x$ are integers. The only possibility is $y = 1/2$ and $x = 1/2 + u$ with $u \in \mathbb{Z}$. Looking at the coordinate on e_2, we get $d = 1 + 2u$. But now, the convexity of C forces the integer points belonging to the line joining the middle of the segment $[O, A]$ to M to belong to E, in particular T'. As $M = T' + u(2T' - \alpha e_1)$ and since there is no line containing more than 2 points of $E \cap \Lambda$, one has $u \leq 1$ i.e. $d \leq 3$. □

Lemma 5. — *Let i, j, k be integers and t a real, $0 < t < 1$, such that $k = ti + (1-t)j$. Then the following holds*

(i) *if $\Delta_i \cap C, \Delta_j \cap C \neq \varnothing$, then one has $c_k \geq tc_i + (1-t)c_j - 2$,*

(ii) *if $\Delta_{\beta i} \cap C, \Delta_{\beta j} \cap C \neq \varnothing$, then one has $c'_{\beta k} \geq tc'_{\beta i} + (1-t)c'_{\beta j} - 2$,*

(iii) $c'_{\beta k} \leq 1 + \dfrac{c_{\beta k} - 1}{\alpha}$.

Proof. — Because of convexity, $\Delta_i \cap C$,$\Delta_j \cap C$ and $\Delta_k \cap C$ are non-empty segments. If l_i, l_j and l_k denote their respective length, one has, once again by convexity, $tl_i + (1-t)l_j \leq l_k$, but one has $l_i - |e_1| \leq c_i|e_1| \leq l_i + |e_1|$ (and the same for j and k), so we get

$$t(c_i - 1) + (1-t)(c_j - 1) \leq c_k + 1,$$

that is the first inequality.

The second one is similar (that is just a question of scale). And the third one is the consequence of an easy counting argument. □

Lemma 6. — *One has*

$$d \leq \frac{2}{\sqrt{\alpha}}|E|^{1/2} + 3 + \frac{4}{3\alpha}.$$

Proof. — We first notice that, because of $|E| \geq 3$, the formula is easily verified in the following cases: $\alpha = 1$ and $d \leq 6$, $\alpha = 2$ or 3 and $d \leq 4$ and $\alpha \geq 4$ and $d \leq 3$ and that except in those cases, which from now on we do not consider anylonger, one has $\alpha(d-3) - 3 \geq 1$. This remark is useful to make easier the forthcoming estimations. If $d \geq 4$, one can write

$$|E| \geq \sum_{k=0}^{\beta d} c_k \geq c_0 + \sum_{k=1}^{[\beta d/2]} (c_0/2 - 2),$$

because of (i) in Lemma 5. This way, we obtain the lower bound

$$|E| \geq c_0 + (c_0/2 - 2)[\beta d/2] \geq n\alpha + 1 + [\beta d/2](n\alpha - 3)/2,$$

with the inequality $c_0 = |\Delta_0 \cap E| \geq n\alpha + 1$; it implies, by virtue of Lemma 4, (ii), that

$$|E| \geq \alpha(d-3) + 1 + [\beta d/2]\frac{\alpha(d-3)-3}{2}.$$

But the preliminary remark of this Lemma ensures that the last fraction is positive. Since $\beta \geq 1$, we get the lower bound for $|E|$:

$$|E| \geq \alpha(d-3) + 1 + \left(\frac{d-1}{2}\right)\left(\frac{\alpha(d-3)-3}{2}\right),$$

which can be rewritten as follows:

$$4|E| \geq \alpha d^2 - 3d + 7 - 9\alpha.$$

It is easy to see that it implies

$$\begin{aligned} d &\leq \frac{3}{2\alpha} + \sqrt{\frac{4|E|}{\alpha} + 9 - \frac{7}{\alpha} + \frac{9}{4\alpha^2}} \\ &\leq \frac{3}{2\alpha} + \frac{2\sqrt{|E|}}{\sqrt{\alpha}} + \sqrt{9 - \frac{7}{\alpha} + \frac{9}{4\alpha^2}} \\ &\leq \frac{3}{2\alpha} + \frac{2\sqrt{|E|}}{\sqrt{\alpha}} + 3 - \frac{7}{6\alpha} + \frac{\sqrt{9/4 - 49/36}}{\alpha}, \end{aligned}$$

which implies the announced result. □

Lemma 7. — *We have*

$$\left(\sum_{i=-d^-+1}^{d^+-1} c_{i\beta}\right) - 4(d^+ + d^-) \geq -39.$$

Proof. — Suppose first that $d^- \geq 1$ (and thus $d^+ \geq 1$). For any positive integer i, we have, in view of Lemma 5, (i),

$$c_{i\beta} \geq \left(1 - \frac{i}{d^+}\right) c_0 + \frac{i}{d^+} c_{\beta d^+} - 2.$$

The same inequality holds, symmetrically, for the $c_{-i\beta}$'s (changing d^+ in d^-). By summing these inequalities, we get

$$\begin{aligned}
\sum_{i=-d^-+1}^{d^+-1} c_{i\beta} &\geq c_0 + \sum_{i=1}^{d^+-1}\left(\left(1-\frac{i}{d^+}\right)c_0 + \frac{i}{d^+}c_{\beta d^+} - 2\right) \\
&\qquad + \sum_{i=1}^{d^--1}\left(\left(1-\frac{i}{d^-}\right)c_0 + \frac{i}{d^-}c_{-\beta d^-} - 2\right) \\
&= c_0 + \left(\frac{c_0 + c_{d+\beta}}{2} - 2\right)(d^+ - 1) + \left(\frac{c_0 + c_{d-\beta}}{2} - 2\right)(d^- - 1) \\
&\geq c_0 + \left(\frac{c_0 + 1}{2} - 2\right)(d^+ + d^- - 2),
\end{aligned}$$

where we have used the fact that $c_{d+\beta}, c_{d-\beta} \geq 1$ because of the non-emptiness of $\Delta_{d+\beta} \cap E$ and $\Delta_{d-\beta} \cap E$. Finally we get

$$\left(\sum_{i=-d^-+1}^{d^+-1} c_{i\beta}\right) - 4(d^+ + d^-) \geq 3 + \frac{c_0 - 11}{2}(d^+ + d^-).$$

Now, we consider two cases. If $c_0 \geq 11$, this is greater than 3. Or else, in view of $c_0 \geq n\alpha + 1 \geq n + 1$, and Lemma 4, (ii), this is $\geq 3 + (n-10)(n+3)$ if $2 \leq n \leq 9$. This expression is minimal for $n = 3$ or 4 and is in these cases equal to -39.

Assume now $d^- = 0$ (recall $d = d^+$), the same inequalities as for the case $d^- \neq 0$ show that our expression is

$$\begin{aligned}
\geq \sum_{i=1}^{d-1} c_{i\beta} - 4d &\geq \left(\frac{c_0 + 1}{2} - 2\right)(d-1) - 4d \\
&= \frac{n-10}{2}(d-1) - 4,
\end{aligned}$$

which is ≥ -4 if $n \geq 10$ and if $n \leq 9$, this is $\geq (n+2)(n-10)/2 - 4 \geq -22$. This completes the proof. □

We are now ready to prove the first proposition of this section. It will be useful for Theorem 3 and algorithmical aspects of our problem but we think that it is an interesting result in itself.

Proposition 4. — *Let C be a compact convex body in $\mathbb{R}^2$ and E denote the set of its integer points. Assume E is not included in a line. Then, for each integer lattice Λ different from $\mathbb{Z}^2$, one has either $E \cap \Lambda$ included in a line, or $|E \cap \Lambda| \leq \frac{2}{3}|E| + 39$.*

Freiman [**F96**] obtained a non-effective version of this result with a "reduction" factor $3/4$ in place of our $2/3$ which is the best possible, as one can see by considering the family depending on an integer parameter n:

$$E_n = C_n \cap \mathbb{Z}^2 = \{(i,j), 0 \leq i \leq n-1, 0 \leq j \leq 2\} \cup \{(n,0), (-1,0)\}$$

(where C_n denotes the convex hull of E_n) and the lattice

$$\Lambda = \mathbb{Z}e_1 + 2\mathbb{Z}e_2,$$

because then $|E_n| = 3n+2$, $|E_n \cap \Lambda| = 2n+2$ and consequently $|E_n \cap \Lambda| = \frac{2}{3}|E_n| + \frac{2}{3}$.

The constant 39 appearing in Proposition 4 seems to be larger than the one one might expect. Once again, the reason is that our computations are rough. Indeed it seems that one could expect a constant very near from 1. This problem of minimizing that constant seems to be open.

Note that the higher dimensional analogue to Proposition 4 is false, contrarily to what is announced in [**C91a**, Lemma 2], as can be seen by considering the following example. In $\mathbb{R}^3$ consider the points $a = (1,0,0), b = (0,-n,0), c = (0,n,0)$ and $d = (-1,0,2)$ for an integer parameter n. Let C'_n be the convex hull of these points

$$E'_n = \mathbb{Z}^3 \cap C'_n = \{(1,0,0),(0,0,1),(-1,0,2),(0,j,0), -n \leq j \leq n\}$$

and $\Lambda = \mathbb{Z}e_1 + \mathbb{Z}e_2 + 2\mathbb{Z}e_3$. One has $E'_n \cap \Lambda = E'_n \setminus \{(0,0,1)\}$ and thus

$$\frac{|E'_n \cap \Lambda|}{|E'_n|} = 1 - \frac{1}{|E'_n|}$$

that tends to 1 as n tends to infinity. At the same time, $E'_n \cap \Lambda$ is 3-dimensional. This shows that no strictly less than 1 analogue to the constant 2/3 exists in dimension 3.

Proof of Proposition 4. — If Λ is not $\mathbb{Z}^2$ then α or β is different from 1, that is at least 2.

First we consider the case where $\alpha \geq 2$. We can write, using Lemma 5 (iii),

$$\begin{aligned} |E \cap \Lambda| = \sum_{k=-d^-}^{d^+} c'_{\beta k} &\leq \sum_{k=-d^-}^{d^+} \left(1 + \frac{c_{\beta k} - 1}{\alpha}\right) \\ &= \left(1 - \frac{1}{\alpha}\right)(1 + d^- + d^+) + \frac{1}{\alpha} \sum_{k=-d^-}^{d^+} c_{\beta k} \\ &\leq \frac{|E|}{\alpha} + (2d+1). \end{aligned} \tag{36}$$

In view of Lemma 6, equation (36) can be rewritten, because of $\alpha \geq 2$,

$$\begin{aligned} |E \cap \Lambda| &\leq \frac{|E|}{2} + 2\left(\sqrt{2}|E|^{1/2} + \frac{11}{3}\right) + 1 \\ &\leq \frac{|E|}{2} + 2\sqrt{2}|E|^{1/2} + \frac{25}{3}, \end{aligned}$$

which is bounded by $\frac{2}{3}|E| + 21$, as one can easily check.

Now we consider the case where $\alpha = 1$. Then one has $\beta \geq 2$ and $c'_{\beta k} = c_{\beta k}$. We have the trivial lower bound

$$|E| \geq \sum_{k=-\beta d^-}^{\beta d^+} c_k. \tag{37}$$

We put for $0 \leq i \leq d^+ - 1$,

$$S_i^+ = \sum_{k=i\beta}^{(i+1)\beta-1} c_k,$$

and, likewise, for $0 \leq i \leq d^- - 1$,

$$S_i^- = \sum_{k=-i\beta}^{-(i+1)\beta+1} c_k;$$

then equation (37) becomes

$$(38) \qquad |E| + c_0 \geq (S_0^+ + \cdots + S_{d^+-1}^+ + c_{d^+\beta}) + (S_0^- + \cdots + S_{d^--1}^- + c_{-d^-\beta}).$$

But, application of Lemma 5, after summation, yields

$$S_i^+ = c_{i\beta} + \sum_{k=i\beta+1}^{(i+1)\beta-1} c_k \geq \frac{\beta+1}{2} c_{i\beta} + \frac{\beta-1}{2} c_{(i+1)\beta} - 2(\beta-1),$$

and, symmetrically,

$$S_i^- \geq \frac{\beta+1}{2} c_{-i\beta} + \frac{\beta-1}{2} c_{-(i+1)\beta} - 2(\beta-1),$$

so equation (38) implies

$$\begin{aligned} |E| + c_0 \quad \geq \quad & \left(\frac{\beta+1}{2} \sum_{i=0}^{d^+-1} c_{i\beta} + \frac{\beta-1}{2} \sum_{i=1}^{d^+} c_{i\beta} - 2(\beta-1)d^+ + c_{d^+\beta} \right) \\ & + \left(\underbrace{\frac{\beta+1}{2} \sum_{i=0}^{d^--1} c_{-i\beta} + \frac{\beta-1}{2} \sum_{i=1}^{d^-} c_{-i\beta}}_{=0 \text{ if } d^-=0} -2(\beta-1)d^- + c_{d^-\beta} \right), \end{aligned}$$

which takes the following simplified form

$$|E| \geq \frac{\beta+1}{2} \sum_{k=-d^-}^{d^+} c_{k\beta} + \frac{\beta-1}{2} \left(\sum_{k=-d^-+1}^{d^+-1} c_{k\beta} - 4(d^+ + d^-) \right).$$

But the first sum is $|E \cap \Lambda|$ and the second ≥ -39 in view of Lemma 7, so we have

$$|E| \geq \frac{\beta+1}{2} |E \cap \Lambda| - 39 \frac{\beta-1}{2}.$$

Thus,

$$|E \cap \Lambda| \leq \frac{2}{\beta+1} |E| + 39 \frac{\beta-1}{\beta+1} \leq \frac{2}{\beta+1} |E| + 39 \leq \frac{2}{3} |E| + 39.$$

□

From now on, we are only interested in E itself (which corresponds to $\Lambda = \mathbb{Z}^2$ or equivalently to $\alpha = \beta = 1$) for which we need two more lemmas. We keep the same notation as above but, for avoiding confusion, we put an index E so that d, n have nothing to do with d_E and n_E.

Lemma 8. — *Let k be an integer such that $1 \leq k \leq d_E$, then*

$$ke_2 + te_1 \in E \textit{ implies } -k+1 \leq t \leq n_E + k - 1.$$

If $-d_E^- \leq k \leq -1$, then

$$ke_2 + te_1 \in E \textit{ implies } k \leq t \leq 2n_E - k + 1.$$

Proof. — Remember that, by construction, $O, n_E e_1, e_2$ belong to E while $-e_1, e_2 - e_1$, $e_2 + (n_E + 1)e_1$ do not belong to E.

Let $k \geq 1$.

If $ke_2 + te_1 \in E$ then, as $O \in E$ too, one would have, by convexity

$$e_2 + \frac{t}{k}e_1 = \frac{1}{k}(ke_2 + te_1) + \frac{k-1}{k}O \in C.$$

Suppose $t \leq -k \leq 0$, then

$$e_2 - e_1 = -\frac{k}{t}\left(e_2 + \frac{t}{k}e_1\right) + \left(1 + \frac{k}{t}\right)e_2 \in E,$$

which is not true. Thus $t \geq -k + 1$.

On the other hand, if $ke_2 + te_1 \in E$, one would have also

$$e_2 + \left(\frac{t + (k-1)n_E}{k}\right)e_1 = \frac{1}{k}(ke_2 + te_1) + \left(\frac{k-1}{k}\right)n_E e_1 \in C.$$

Suppose $t \geq n_E + k$, then

$$e_2 + (n_E + 1)e_1 =$$
$$\left(\frac{(n_E+1)k}{t + (k-1)n_E}\right)\left(e_2 + \left(\frac{t + (k-1)n_E}{k}\right)e_1\right) + \left(\frac{t - n_E - k}{t + (k-1)n_E}\right)e_2 \in E,$$

which is not true. Thus $t \leq n_E + k - 1$.

Now, let $k \leq -1$.

If $ke_2 + te_1 \in E$, then, since $e_2 \in E$,

$$\left(\frac{t}{1-k}\right)e_1 = \left(\frac{1}{1-k}\right)(ke_2 + te_1) + \left(\frac{-k}{1-k}\right)e_2 \in C.$$

Suppose $t \leq k - 1$, then

$$-e_1 = \left(\frac{k-1}{t}\right)\left(\frac{t}{1-k}e_1\right) + \left(\frac{t+1-k}{t}\right)O \in E,$$

which is false. Thus $t \geq k$.

Suppose $ke_2 + te_1 \in E$. It is known, by construction, that there is some point $d_E e_2 + xe_1 \in E$ and that, as previously seen, $x \geq 1 - d_E$. But then

$$\left(\frac{d_E t - xk}{d_E - k}\right)e_1 = \left(\frac{d_E}{d_E - k}\right)(ke_2 + te_1) + \left(\frac{-k}{d_E - k}\right)(d_E e_2 + xe_1) \in C.$$

Suppose $t \geq 2n_E + 2 - k$, since

$$\frac{d_E t - xk}{d_E - k} \geq \frac{d_E(2n_E + 2 - k) - k(1 - d_E)}{d_E - k} = \frac{2n_E d_E + 2d_E - k}{d_E - k} \geq n_E + 1,$$

the point $(n_E + 1)e_1$ is located on the segment joining O to $\left(\frac{d_E t - xk}{d_E - k}\right) e_1$ and is consequently in E, which is not true. Thus $t \leq 2n_E + 1 - k$. □

Lemma 9. — *One has*

$$|E| \geq n_E d_E/3.$$

Proof. — As in the previous proofs, one has:

$$\begin{aligned} |E| = c_0 + c_{d_E} + \sum_{i=1}^{d_E - 1} c_i &\geq (c_0 + c_{d_E})(d_E + 1)/2 - 2(d_E - 1) \\ &\geq (n_E + 2)(d_E + 1)/2 - 2(d_E - 1). \end{aligned}$$

Then $(|E| - n_E d_E/3) \geq 0$ follows from $(n_E - 6)d_E + 3n_E + 18 \geq 0$. For $n_E \geq 6$ this is trivially true and one checks that for $n_E \leq 6$, using Lemma 4,

$$(n_E - 6)d_E + 3n_E + 18 \geq n_E^2 \geq 0.$$

□

Now, we prove the second geometric lemma, which will be the key result for obtaining Theorem 2. We have here to remember that A_0 can be different from O.

Proposition 5. — *Let C be a compact convex body containing O, and E denote $C \cap \mathbb{Z}^2$. Then there exists a unimodular linear application ϕ and two integers l, m such that*

$$\phi(E) \subset P_{l,m},$$

with $v = (2l + 1)(2m + 1) \leq 345|E|$.

Proof. — Using the construction described before with $\alpha = \beta = 1$, we find a point A_0 and two integer vectors e_1, e_2 such that

$$\begin{aligned} C &\subset \{A_0 + \{-d_E, \dots, 2n_E + d_E + 1)\}e_1 + \{-d_E, \dots, d_E\}e_2\} \\ &\subset \{A_0 + \{-(n_E + 3), \dots, 3n_E + 4)\}e_1 + \{-d_E, \dots, d_E\}e_2\} \end{aligned}$$

in view of Lemma 8 and $d_E \leq n_E + 3$ (Lemma 4).

Recall that (ϵ_1, ϵ_2) is the canonical basis. Let ϕ be the linear transformation sending the $\mathbb{Z}^2$-basis $(e_i)_{i=1,2}$ onto the $\mathbb{Z}^2$-basis $(\epsilon_i)_{i=1,2}$. We have $\det \phi = \pm 1$ (ϕ is unimodular) and

$$\phi(C) \subset \{\phi(A_0) + \{-(n_E + 3), \dots, (3n_E + 4)\}\epsilon_1 + \{-d_E, \dots, d_E\}\epsilon_2\}.$$

But $\phi(O) = O \in C$ implies the existence of some r, s such that

$$O = \phi(A_0) + r\epsilon_1 + s\epsilon_2,$$

with $-(n_E + 3) \leq r \leq 3n_E + 4, |s| \leq d_E$. This fact shows that $\phi(C) \subset P_{4n_E+7, 2d_E}$, whose volume is $v = (8n_E + 15)(4d_E + 1) \leq 115 n_E d_E$. But, by Lemma 9, $|E| \geq n_E d_E/3 \geq v/345$, which ends the proof. □

3. Proof of Theorem 1

We assume without loss of generality that

$$2l_1 + 1 \geq v^{1/2} \geq 2l_2 + 1.$$

Let us remember that $v \geq k_4$. Define $m = |A|$ and write $m_0 = k_1 v^{2/3} \log^{1/3} v$.

The trivial orthogonality relations for e type functions easily yield

$$J(b) = 2^m \int_{[0,1]^2} \phi(\alpha) e(-b.\alpha) d\alpha = 2^m \int_{[-1/2,1/2]^2} \phi(\alpha) e(-b.\alpha) d\alpha,$$

with

$$\phi_j(\alpha) = \frac{1 + e(a_j.\alpha)}{2}$$

and $\phi = \prod_{j=1}^m \phi_j$. In fact, we investigate $I = I(b) = J(b)/2^m$, by splitting the domain of integration into two parts, the major and minor arcs, corresponding respectively to the domains

$$K_0 = [-1/4l_1, 1/4l_1] \times [-1/4l_2, 1/4l_2]$$

and $K_1 = [-1/2, 1/2]^2 \setminus K_0$. The corresponding integrals are denoted respectively I_0 and I_1.

3.1. The error term. — We have $|I_1| \leq \int_{K_1} |\phi(\alpha)| d\alpha \leq \mu(K_1) \sup_{\alpha \in K_1} |\phi(\alpha)|$. Applying (8) we get

$$|I_1| \leq \mu(K_1) \exp\left(-\frac{\pi^2}{2} \inf_{\alpha \in K_1} \sum_{j=1}^m ||\alpha.a_j||^2 \right).$$

Our main aim in this section is to find a uniform lower bound on K_1 for the sum $\sum_{j=1}^m ||\alpha.a_j||^2$ appearing in the exponential. For this, let us use a Farey dissection of order Q of $[-1/2, 1/2] \setminus [-1/4l_1, 1/4l_1]$: we write (modulo 1) each α_1 of this interval as $p/q + z$ with

$$\gcd(p, q) = 1, 0 < q \leq Q \text{ and } |z| \leq 1/qQ.$$

Here we choose $Q = [k_{12} v/m]$, with

$$k_{12} = 69.$$

We notice that

$$\frac{2l_1 + 1}{Q} \geq \frac{v^{1/2}}{k_{12} v/m} \geq \frac{m_0}{k_{12} v^{1/2}} \geq \frac{k_1}{k_{12}} v^{1/6} \log^{1/3} v > 2, \tag{39}$$

because $v \geq k_4$.

Now, we distinguish different cases.

3.1.1. $|z| \geq 1/4ql_1$. — Using Proposition 1 with $P = Q$, $k = 2l_1$ which is $\geq Q$, $a = \alpha_1$, $b = -\alpha_1 l_1 + \alpha_2 n_2$ and $n = -l_1$, we get for each n_2 subject to $-l_2 \leq n_2 \leq l_2$:

$$|\{-l_1 \leq n_1 \leq l_1 : ||\alpha.n|| \leq Q^{-1}\}| \leq 3(8l_1 Q^{-1} + 1).$$

Thus we get, after summation on n_2,

$$|\{n \in P_{l_1,l_2} : ||\alpha.n|| \leq Q^{-1}\}| \leq 3(8l_1 Q^{-1} + 1)(2l_2 + 1),$$

and this is

$$\leq 3\left(\frac{8l_1+4}{Q} + \frac{2l_1+1}{2Q}\right)(2l_2+1) = 13.5vQ^{-1},$$

in view of (39).

3.1.2. $|z| < 1/4ql_1$. — Write $\alpha_2 = (h+\theta)/q$ where $h \in \mathbb{Z}, 0 \leq \theta < 1$ and $\theta = p'/q' + z'$ by using a Farey dissection of order $4l_2$ (therefore $q' \leq 4l_2$ and $|z'| \leq 1/4q'l_2$). We have

$$\begin{aligned} \alpha.n &= \left(\frac{p}{q} + z\right) n_1 + \left(\frac{h}{q} + \frac{p'}{qq'} + \frac{z'}{q}\right) n_2 \\ &= \frac{q'(pn_1 + hn_2) + p'n_2}{qq'} + zn_1 + \frac{z'n_2}{q}. \end{aligned}$$

If $D_{a,b}$ denotes the set $\{aq, aq+1, \dots, (a+1)q - 1\} \times \{bq', bq'+1, \dots, (b+1)q' - 1\}$, define Ψ the application

$$\begin{aligned} \Psi : \quad D_{a,b} &\longrightarrow \quad \mathbb{Z}/qq'\mathbb{Z} \\ (x,y) &\longmapsto \quad q'(px + hy) + p'y. \end{aligned}$$

It is easy to check that Ψ is bijective (injectivity is just a trivial consequence of $\gcd(p', q') = 1$).

Finally, equation $||\alpha.n|| \leq Q^{-1}$ implies

$$\left|\left|\frac{\Psi(n)}{qq'}\right|\right| \leq Q^{-1} + |z|n_1 + \frac{|z'|n_2}{q} \leq Q^{-1} + |z|l_1 + \frac{1}{4qq'}, \tag{40}$$

in view of $|zn_1| \leq |z|l_1$ and $|z'n_2/q| \leq 1/4qq'$.

3.1.2.1. *Case 1:* $Q^{-1} \geq |z|l_1, 1/4qq'$. — We have

$$Q^{-1} + |z|l_1 + \frac{1}{4qq'} \leq 3Q^{-1},$$

and since Ψ bijective,

$$\begin{aligned} |\{n \in P_{l_1,l_2} : ||\alpha.n|| \leq Q^{-1}\}| &\leq |\{n \in P_{l_1,l_2} : ||\Psi(n)/qq'|| \leq 3Q^{-1}\}| \\ &\leq (1 + 6Q^{-1}qq')([2l_1/q] + 1)([2l_2/q'] + 1), \end{aligned}$$

because $[2l_1/q] + 1$ is the maximal number of integers between $-l_1$ and l_1 having same residue modulo q.

Now, $1 + 6Q^{-1}qq' \leq 10Q^{-1}qq'$ by hypothesis of case 1. One has

$$[2l_1/q] + 1 \leq \frac{2l_1+1}{q} + 1 \leq \frac{3}{2}\left(\frac{2l_1+1}{q}\right),$$

in view of (39). And $[2l_2/q'] + 1 \leq 3(2l_2+1)/q'$ because $q' \leq 4l_2$. Finally,

$$|\{n \in P_{l_1,l_2} : ||\alpha.n|| \leq Q^{-1}\}| \leq 10Q^{-1}qq'(3(2l_1+1)/2q)(3(2l_2+1)/q') \leq 45vQ^{-1}.$$

3.1.2.2. *Case 2: $Q^{-1}, |z|l_1 < 1/4qq'$.* — In this case, a solution of $||\alpha.n|| \leq Q^{-1}$ verifies $||\Psi(n)/qq'|| < 1/qq'$, but this implies $\Psi(n) = 0 \bmod qq'$, that is

$$q'(pn_1 + hn_2) + p'n_2 \equiv 0 \bmod qq'.$$

This implies that $q'|n_2$, so that, when $q' \neq 1$, n belongs to some lattice different from $\mathbb{Z}^2$ (namely $\mathbb{Z}\epsilon_1 + \mathbb{Z}q'\epsilon_2$).

If $q' = 1$, one has $pn_1 + (h+p')n_2 \equiv 0 \bmod q$, which is the equation of a lattice different from $\mathbb{Z}^2$ as soon as $q \neq 1$, but that is the case because if $q = q' = 1$, one has $p = p' = 0$ (and then $h = 0$), and $|z| < 1/4l_1, |z'| \leq 1/4l_2$. Therefore $|\alpha_i| < 1/4l_i$ for $i = 1, 2$, which shows that, in this case, α is not in K_1.

Consequently, we can bound the number of solutions in this case, using hypothesis (3):

$$|\{n \in P_{l_1,l_2} : ||\alpha.n|| \leq Q^{-1}\}| \leq m - k_2 v^{2/3} \log^{1/3} v.$$

3.1.2.3. *Case 3: $|z|l_1 \geq Q^{-1}, 1/4qq'$.* — If the integer vector n satisfies

$$||\alpha.n|| \leq Q^{-1}, \tag{41}$$

it satisfies a fortiori

$$||\Psi(n)/qq'|| \leq 3|z|l_1 \tag{42}$$

and so the number of couples of residues (x_0, y_0), modulo q and q' respectively, solutions to (42), is less than $1 + 6|z|l_1qq'$.

Let us now give an upper bound for the number of solutions of (41) with n_1 restricted to be equal to some x_0 modulo q and n_2 fixed. The equation becomes

$$\left|\left|\frac{pn_1}{q} + z(x_0 + qt) + \alpha_2 n_2\right|\right| \leq Q^{-1};$$

this is of the form

$$||\eta + zqt|| \leq Q^{-1},$$

for which we can apply Lemma 1 with $a = zq, b = \eta, \epsilon = Q^{-1} \leq k_{12}^{-1} < 1/6, k = [2l_1/q] + 1$. We have $k|a| \leq |z|q(1 + 2l_1/q) = |z|q + 2l_1|z| < Q^{-1} + 1/2q \leq Q^{-1} + 1/2$, by hypothesis of section 3.1.2 and thus $(1 - k|a|)/2 > \epsilon$ is verified. We get the upper bound $1 + [2/Q|z|q]$. Consequently, if now, n_1 and n_2 are restricted to be constant modulo q and q' respectively, the number of solutions of (41) is

$$\leq \left(1 + \left[\frac{2l_2}{q'}\right]\right)\left(1 + \frac{2}{Q|z|q}\right).$$

Finally, the total number of possible solutions to (41) is bounded above by

$$(1 + 6|z|l_1qq')\left(1 + \left[\frac{2l_2}{q'}\right]\right)\left(1 + \frac{2}{Q|z|q}\right) \leq 45vQ^{-1},$$

by using $|z|l_1qq' \geq 1/4$, $|z| \leq 1/qQ$ and $q' \leq 4l_2$.

3.1.3. *Conclusion.* — In the cases of sections 3.1.1 and 3.1.2 cases 1 and 3, the total number of solution to (41) is bounded by

$$45vQ^{-1} \leq \frac{45}{k_{12}-1}m.$$

Consequently,

$$|\{n \in A, ||\alpha.n|| \geq Q^{-1}\}| \geq \frac{k_{12}-46}{k_{12}-1}m.$$

Thus we get

$$\sum_{j=1}^{m} ||\alpha.a_j||^2 \geq \frac{k_{12}-46}{k_{12}-1}mQ^{-2} \geq \frac{k_{12}-46}{(k_{12}-1)k_{12}^2}\frac{m^3}{v^2} \geq \frac{(k_{12}-46)k_1^3}{(k_{12}-1)k_{12}^2}\log v \geq 0.75\log v.$$

In the case of section 3.1.2, case 2, we have

$$|\{n \in A, ||\alpha.n|| \geq Q^{-1}\}| \geq k_2 v^{2/3}\log v,$$

thus

$$\sum_{j=1}^{m} ||\alpha.a_j||^2 \geq \left(\frac{m}{k_{12}v}\right)^2 k_2 v^{2/3}\log^{1/3} v \geq \frac{k_1^2 k_2}{k_{12}^2}\log v \geq 0.75\log v.$$

Finally,

$$|I_1| \leq \mu(K_1)\exp\left(-\frac{\pi^2}{2}\inf_{\alpha\in K_1}\sum_{j=1}^{m}||\alpha.a_j||^2\right) \leq \mu(K_1)/v^3.$$

3.2. The major part. — Here, we have to investigate $I_0 = \int_{K_0}\phi(\alpha)e(-b.\alpha)d\alpha$. Let us denote

$$K = \{\alpha \in K_0 : V(\alpha) \leq 0.75\log v\}$$

and $K' = K_0 \setminus K$. The contribution of K' can be evaluated as follows

$$\begin{aligned}\left|\int_{K'}\phi(\alpha)e(-b.\alpha)d\alpha\right| &\leq \int_{K'}\prod_{j=1}^{m}|\phi_j(\alpha)|d\alpha \leq \mu(K')\exp\left(-\pi^2\sum_{j=1}^{m}||\alpha.a_j||^2/2\right)\\ &\leq \mu(K')\exp\left(-\pi^2 V(\alpha)/2\right) \leq \mu(K')/v^{3\pi^2/8},\end{aligned}$$

in view of the definition of V and because on K_0, $|\alpha_i a_{j,i}| \leq 1/4$ so that $||\alpha.a_j|| = |\alpha.a_j| \leq 1/2$.

Now, rewrite $\phi(\alpha)e(-b.\alpha)$ as follows

$$\phi(\alpha)e(-b.\alpha) = e((M-b).\alpha)\prod_{j=1}^{m}\cos(\pi a_j.\alpha),$$

but for $|\pi a_j.\alpha| \leq \pi/2$, one can write, in view of (9),

$$\cos(\pi a_j.\alpha) = \exp(-\pi^2(a_j.\alpha)^2/2)(1-g(a_j.\alpha)),$$

with

$$0 \leq g(a_j.\alpha) \leq (2\pi a_j.\alpha/\pi)^4 \leq 1.$$

Finally the major part is

$$\int_K \phi(\alpha)e(-b.\alpha)d\alpha = \int_K \exp\left(-\frac{\pi^2}{2}\sum_{j=1}^{m}(a_j.\alpha)^2 + 2i\pi(M-b).\alpha\right)(1-R(\alpha))d\alpha,$$

where $R(\alpha) = 1 - \prod_{j=1}^{m}(1-g(a_j.\alpha))$. One can write this integral $A_0 - A_1 - A_2$ by splitting it into three parts

$$\begin{aligned} A_0 &= \int_{\mathbb{R}^2} \exp\left(-\frac{\pi^2}{2}V(\alpha) + 2i\pi(M-b).\alpha\right)d\alpha, \\ A_1 &= \int_{\mathbb{R}^2\setminus K} \exp\left(-\frac{\pi^2}{2}V(\alpha) + 2i\pi(M-b).\alpha\right)d\alpha, \\ A_2 &= \int_K R(\alpha)\exp\left(-\frac{\pi^2}{2}V(\alpha) + 2i\pi(M-b).\alpha\right)d\alpha. \end{aligned}$$

Let us write $d_i = M_i - b_i$ and investigate these three integrals. By the change of variables

$$x = \frac{\pi}{V_1}(V_1^2\alpha_1 + V_{12}\alpha_2), y = \pi\frac{\sqrt{\det V}}{V_1}\alpha_2, \tag{43}$$

we get

$$\begin{aligned} A_0 &= \frac{1}{\pi^2\sqrt{\det V}}\int_{\mathbb{R}}\exp\left(-\frac{x^2}{2} + i\frac{2d_1x}{V_1}\right)dx\int_{\mathbb{R}}\exp\left(-\frac{y^2}{2} + i\frac{2(d_2V_1^2 - d_1V_{12})y}{V_1\sqrt{\det V}}\right)dy \\ &= \frac{1}{\pi^2\sqrt{\det V}}\sqrt{2\pi}\exp\left(-\frac{1}{2}\left(2\frac{d_1}{V_1}\right)^2\right)\sqrt{2\pi}\exp\left(-\frac{1}{2}\left(2\frac{d_2V_1^2 - d_1V_{12}}{V_1\sqrt{\det V}}\right)^2\right) \\ &= \frac{2\exp(-2q_{V^{-1}}(M-b))}{\pi\sqrt{\det V}}. \end{aligned}$$

An upper bound for $|A_1|$ can be achieved by noticing that if $\alpha \notin K$ then $V(\alpha) \geq 0.75\log v$. Suppose $\alpha \notin \mathbb{Z}^2 + K_0$. We have

$$V(\alpha) = \sum_{i=1}^{m}(\alpha.a_j)^2 \geq \sum_{j=1}^{m}||\alpha.a_j||^2.$$

Since the last sum is not changed by an integral translation of α, the problem is reduced to the study of this sum for $\alpha \in K_1$. This has been done in the preceding section and thus we get the lower bound $0.75\log v$. If now $\alpha \in (\mathbb{Z}^2 \setminus \{(0,0)\}) + K_0$, it can be written as $\alpha = h + \epsilon$ with $h \in \mathbb{Z}^2 \setminus \{(0,0)\}$ and $\epsilon \in K_0$. Since $|\epsilon_i| \leq 1/4l_i$, $|\epsilon.a_j| \leq 1/2$ and in view of hypothesis (3) at least $k_2v^{2/3}\log^{1/3}v$ elements a_j of A verify $h.a_j \neq 0$ (cf. hypothesis (28)) thus $(\alpha.a_j)^2 \geq 1/4$ for these values and we obtain

$$V(\alpha) \geq \frac{k_2}{4}v^{2/3}\log^{1/3}v \geq 0.75\log v.$$

This ends the proof of the lower bound of $V(\alpha)$ on the complementary of K. Now, the change of variables (43) and a polar change of variables produce

$$|A_1| \leq \frac{1}{\pi^2\sqrt{\det V}} \int_{\pi\sqrt{3\log v/4}}^{+\infty} r\exp(-r^2/2)dr \int_0^{2\pi} d\theta,$$

and finally

$$|A_1| \leq 2/(\pi\sqrt{\det V}v^{3\pi^2/8}).$$

Now, we consider A_2. We have, in view of inequality (10),

$$|R(\alpha)| = |1 - \prod_{j=1}^{m}(1 - g(a_j.\alpha))| \leq \sum_{j=1}^{m} g(a_j.\alpha) \leq 16\sum_{j=1}^{m}(a_j.\alpha)^4,$$

the last inequality being due to (9). Here is the place where we need Proposition 2. We get

$$\begin{aligned}|A_2| &\leq 16\int_K \left(\sum_{j=1}^{m}(a_j.\alpha)^4\right) \exp\left(-\frac{\pi^2}{2}V(\alpha)\right) d\alpha \\ &\leq \frac{16}{k_{11}\log v}\int_{\mathbb{R}^2} V(\alpha)^2 \exp\left(-\frac{\pi^2}{2}V(\alpha)\right) d\alpha.\end{aligned}$$

In the same way as above we obtain finally

$$|A_2| \leq \frac{256}{k_{11}\pi^5\sqrt{\det V}\log v} \leq \frac{750}{\log v\sqrt{\det V}}.$$

3.3. Conclusion. — The dominant term is

$$A_0 = \frac{2}{\pi\sqrt{\det V}}\exp(-2q_{V^{-1}}(M - b)).$$

The error term is bounded from above by

$$\frac{\mu(K_1)}{v^3} + \frac{\mu(K')}{v^{3\pi^2/8}} + \frac{2}{\pi\sqrt{\det V}v^{3\pi^2/8}} + \frac{750}{\log v\sqrt{\det V}} \leq \frac{800}{\log v\sqrt{\det V}},$$

using $|\det V| \leq m^2v^2/4 \leq v^4/4$. It is readily seen that if $q_{V^{-1}}(M-b) \leq k_3\log\log v - 4$ then the main term A_0 is $\geq 1800/\sqrt{\det V}\log^{2k_3} v$. At the same time the error term is

$$\leq \frac{800}{\sqrt{\det V}\log v} = o(A_0),$$

(with a constant 1) thanks to our choice for k_3. This concludes the proof of Theorem 1.

4. Proof of Theorems 2 and 3

4.1. Theorem 2. — By Proposition 5, we can find a linear application ϕ sending C onto P_{l_1,l_2} for some l_1, l_2 with $v \leq 345|E|$. We can assume that $v \geq |E|$ with no loss of generality. Consequently,

$$\begin{aligned} |\phi(A)| = |A| \geq k_5|E|^{2/3}\log^{1/3}|E| &\geq k_5\left(\frac{v}{345}\right)^{2/3}\log^{1/3}\left(\frac{v}{345}\right) \\ &\geq \frac{k_5}{100}v^{2/3}\log^{1/3}v = k_1v^{2/3}\log^{1/3}v. \end{aligned}$$

Since the linear transformation ϕ sends lattices onto lattices we have

$$|\phi(A)\setminus(\phi(A)\cap\Gamma)| = |A\setminus(A\cap\phi^{-1}(\Gamma))| \geq k_6|E|^{2/3}\log^{1/3}|E|,$$

by hypothesis (6). As above this is

$$\geq \frac{k_6}{100}v^{2/3}\log^{1/3}v = k_2v^{2/3}\log^{1/3}v.$$

Applying Theorem 1 to the set $\phi(A)$, the asymptotic formula (4) is changed in (J' stands for the number of solutions to the boolean equation induced by $\phi(A)$)

$$J(b) = J'(\phi(b)) \sim \frac{2^{m+1}}{\pi\sqrt{\det W}}\exp\{-2q_{W^{-1}}(\phi(M)-\phi(b))\},$$

for any b such that $q_{W^{-1}}(\phi(M)-\phi(b)) \leq k_3\log\log v - 4$, where W is the matrix obtained with the $\phi(a_j)$'s instead of the a_j's that is to say $W = \phi V\phi^t$. Thus $\det W = \det^2\phi\ \det V = \det V$, $q_{W^{-1}}(\phi(M)-\phi(b)) = q_{V^{-1}}(M-b)$ and we can take $k_7 = k_3$ due to $\log\log|E| \leq \log\log v$. □

From Theorem 2 we deduce the following result.

Corollary 1. — *Let C be a compact convex set in $\mathbb{R}^2$ containing O, Λ be an integer lattice and $E = C\cap\Lambda$. Let A be a subset of E. Assume*

$$|A| \geq k_5|E|^{2/3}\log^{1/3}|E|$$

and that for each Γ sub-lattice of Λ different from $\mathbb{Z}^2$, we have

$$|A\setminus A\cap\Gamma| \geq k_6|E|^{2/3}\log^{1/3}|E|. \tag{44}$$

Then we have the following asymptotic equivalent (when $|E| \to +\infty$)

$$J(b) \sim \frac{2^{m+1}\ \mathit{Vol}\,\Lambda}{\pi\sqrt{\det V}}\exp\{-2q_{V^{-1}}(M-b)\}, \tag{45}$$

provided that $q_{V^{-1}}(M-b) \leq k_7\log\log|E| - 4$.

Proof of the Corollary. — Take a basis of Λ and Ψ a linear application sending this basis onto the canonical basis of $\mathbb{Z}^2$. If $A' = \Psi(A), E' = \Psi(C)\cap\mathbb{Z}^2 = \Psi(E)$ and since the sub-lattices of Λ are sent onto integer lattices, we can apply Theorem 2 to $\Psi(C)$ and $\Psi(A)$. With the same computation as above, we get the asymptotic equivalent (45), the factor Vol Λ being due to the formula of change of basis for quadratic forms and $|\det\Psi| = \text{Vol}\ \Lambda$. □

4.2. Theorem 3. — Either condition (6) is fulfilled by A and we are done by applying Theorem 2 (notice $k_8 \geq k_5$) or there is an integral lattice Γ_1 such that

$$|A \cap \Gamma_1| \geq |A| - k_6 |E|^{2/3} \log^{1/3} |E|. \tag{46}$$

Write $\Gamma_0 = \mathbb{Z}^2$, $A_i = A \cap \Gamma_i$ and $E_i = E \cap \Gamma_i$. Since A_1 is not contained in a line in view of hypothesis (7) ($|A_1| \geq \left(1 - \frac{k_6}{k_8}\right) |A|$ and $k_9 < 1 - k_6/k_8$) we have, by Proposition 4, $|E_1| \leq \frac{2}{3}|E| + 39 \leq 0.7|E_0|$ (the smallest possible value of $|E|$ allows to write this). Therefore we get

$$|A_1| \geq k_8 |E_1|^{2/3} \log^{1/3} |E_1| \tag{47}$$

in view of equation (46) and $(0.7)^{2/3} k_8 \leq k_8 - k_6$.

Now either A_1 verifies condition (44) of Corollary 1 and we stop here the process, or there exists a lattice $\Gamma_2 \subset \Gamma_1$ violating (44).

Let us show that, more generally, we can construct a decreasing finite sequence of lattices $(\Gamma_i)_{1 \leq i \leq p}$: assume we have already built $\Gamma_1, \ldots, \Gamma_i$ and $A_1, \ldots, A_i$. If condition (44) is fulfilled for $A_i \subset E_i$ then we stop the process else we find a lattice $\Gamma_{i+1} \subset \Gamma_i$ violating (44).

Lemma 10. — *We have for each* i

$$|A_i| \geq k_8 |E_i|^{2/3} \log^{1/3} |E_i| \textit{ and } |E_i| \geq |A_i| > |A|/2.$$

Proof. — For $i = 1$, equation (47), $k_8 > 2k_6$ and (46) prove the result. Assume that the result is true for $1, 2, \ldots, i$ and that we have built Γ_{i+1}, E_{i+1} and A_{i+1}. One has

$$|A_{i+1}| \geq |A_i| - k_6 |E_i|^{2/3} \log^{1/3} |E_i| \geq$$
$$(k_8 - k_6)|E_i|^{2/3} \log^{1/3} |E_i| \geq k_8 |E_{i+1}|^{2/3} \log^{1/3} |E_{i+1}|$$

in view of Proposition 4 (here we used the fact that $|A_{i+1}| \geq (1 - k_6/k_8)|A_i| \geq \frac{1}{2}(1 - k_6/k_8)|A| \geq k_9|A|$ which implies first that E_{i+1} is not included in a line (in view of (7)) and, second, that $|E_i|$ is large enough). Now

$$\begin{aligned} |A_{i+1}| &\geq |A| - k_6 \sum_{j=0}^{i} |E_j|^{2/3} \log^{1/3} |E_j| \\ &\geq |A| - k_6 \sum_{j=0}^{\infty} (0.7^j |E|)^{2/3} \log^{1/3} |E| \\ &\geq |A| - 5k_6 |E|^{2/3} \log^{1/3} |E| > |A|/2 \end{aligned}$$

due to the definition of k_6 and k_8. □

This Lemma shows that the process is well defined. As it is clearly finite (since $(|E_i|)_{1 \leq i \leq p}$ is a strictly decreasing sequence and E_i is never included in a line in view of hypothesis (7) and Lemma 10) and at the end we have a lattice $\Lambda_0 = \Gamma_p$, $A_p = A \cap \Gamma_p$ and $E_p = E \cap \Gamma_p$ such that condition (44) and the cardinality condition are fulfilled. Thus we can apply the Corollary to Theorem 2 which gives the result.

References

[AF88] Alon N. and Freiman G. A., *On sums of subsets of a set of integers*, Combinatorica, **8 (4)**, 1988, 297–306.

[C] Cassels J. W. S., *An introduction to the geometry of numbers*, Springer Verlag, 1971.

[C91a] Chaimovich M., *On solving dense n-dimensional subset sum problems*, Congressus Numerantium, **84**, 1991, 41–49.

[C91b] Chaimovich M., *Analytical methods of number theory in integer programming*, PhD, University of Tel-Aviv, 1991.

[CFG89] Chaimovich M., Freiman G. A. and Galil Z., *Solving dense subset sum problem by using analytical number theory*, J. of Complexity, **5**, 1989, 271–282.

[EF90] Erdős P. and Freiman G. A., *On two additive problems*, J. Number Theory, **34**, 1990, 1–12.

[F80] Freiman G. A., *An analytical method of analysis of linear boolean equations*, Ann. New-York Acad. Sci., **337**, 1980, 97–102.

[F93] Freiman G. A., *New analytical results in subset sum problem*, Discrete Math., **114**, 1993, 205–217. For erratum, see Discrete Math., **126**, 1994, 447.

[F96] Freiman G. A., *On solvability of a system of two boolean linear equations*, Number Theory: New-York Seminar 1991-1995, Springer-Verlag, 1996, 135–150.

[HW] Hardy G. W. and Wright E. M., *An introduction to the theory of numbers*, 5th ed., Oxford University Press, 1979.

A. Plagne, Algorithmique Arithmétique Expérimentale, CNRS UMR 9936,
Université Bordeaux I, 351 cours de la Libération, 33405 Talence Cedex, FRANCE
E-mail : plagne@math.u-bordeaux.fr