Available online at www.sciencedirect.com

ScienceDirect

Indagationes Mathematicae 25 (2014) 632–645

indagationes mathematicae

www.elsevier.com/locate/indag

Geodesic continued fractions and LLL

Frits Beukers

Utrecht University, Department of Mathematics, Utrecht, Netherlands

Abstract

We discuss a proposal for a continued fraction-like algorithm to determine simultaneous rational approximations to d real numbers $\alpha_1, \ldots, \alpha_d$. It combines an algorithm of Hermite and Lagarias with ideas from LLL-reduction. We dynamically LLL-reduce a quadratic form with parameter t as $t \downarrow 0$. Suggestions in this direction have been made several times over in the literature, e.g. Chevallier (2013) [4] or Bosma and Smeets (2013) [2]. The new idea in this paper is that checking the LLL-conditions consists of solving linear equations in t.

Keywords: Multidimensional continued fraction; Minkowski reduction; LLL-reduction

1. Introduction

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ and suppose that not all α_i are rational. Dirichlet showed that there exist infinitely many $(\mathbf{p}, q) = (p_1, p_2, \ldots, p_d, q) \in \mathbb{Z}^{d+1}$ with gcd one and $q > 0$ such that

$$\max_{i=1}^{d} |p_i - q\alpha_i| \leq q^{-1/d}. \tag{D}$$

The proof is a simple application of the pigeon-hole principle.

A more difficult problem is to actually compute these simultaneous approximations in an efficient way. The Jacobi–Perron algorithm and related versions of it (modified Perron, Brun, Selmer) seem to yield upper bounds like $c \cdot q^{-\delta}$ but with $0 < \delta < 1/d$ for general $\boldsymbol{\alpha} \in \mathbb{R}^d$. So they are not expected to be very good.

In a letter to Jacobi, Hermite explained another idea to construct good simultaneous approximations, [8, p. 106]. See also [12, p. xii, xiii]. Choose $t > 0$ and consider the quadratic

E-mail address: f.beukers@uu.nl.

http://dx.doi.org/10.1016/j.indag.2014.04.003

form

$$Q_t = (x_1 - \alpha_1 y)^2 + \cdots + (x_d - \alpha_d y)^2 + ty^2$$

in $x_1, \ldots, x_d, y$. Choose integers $p_1, \ldots, p_d, q$ as arguments which minimize this form. Then Hermite was able to show that

$$|\mathbf{p} - \boldsymbol{\alpha} q| \leq \gamma_d q^{-1/d}$$

where γ_d is a number depending only on d. The difference here is that we now work with the Euclidean norm $|\cdot|$ on $\mathbb{R}^d$ rather than the supremum norm in Dirichlet's theorem, but we will allow for that. For a proof see Proposition 4.2 in this paper. All that is required now, is a reduction algorithm that enables one to find the minimizing set of integers.

In 1994 Jeff Lagarias, in [10], took up this idea again and proposed an algorithm which consists in decreasing t to 0 and along the way performs coordinate changes so that the form remains Minkowski reduced (see Section 3 for a definition). The result is an algorithm of the type sketched on p. 12. A similar elaboration in the case $d = 1$ can already be found in a paper by Humbert, [9]. The result is an algorithm that produces good simultaneous approximations like (D). In [10] Lagarias gives an analysis of this algorithm. For example, it finds best approximations in the Euclidean norm sense. That is, it finds $q \in \mathbb{Z}_{>0}$ such that $\|q\boldsymbol{\alpha}\| \leq \|q'\boldsymbol{\alpha}\|$ for all integers q' with $0 < q' < q$. However, it is not guaranteed that all of them are found.

A nice feature of Lagarias' algorithm is that the Minkowski reducedness conditions are linear in t, which makes the check and update process easy. The disadvantage is that the number of conditions grows prohibitively large as d increases. Already for $d = 7$ about 90,000 conditions are needed.

This problem might be circumvented by the use of LLL-reduction instead of Minkowski reduction (again see Section 3 for a definition). Since the LLL-algorithm gives suboptimal results, one cannot expect to find guaranteed best approximations. If one is willing to accept this, another potential problem is that the LLL-reduction conditions are non-linear in the coefficients of Q_t, thus making their verification difficult. However, the main contribution of this paper is the observation that the conditions are still linear in t. The algorithm we propose in Section 4 is not very surprising, but its feasibility is based on the linearity in t of the LLL-conditions. Another feature of the algorithm presented here is that it is dynamic in the sense of a true continued fraction algorithm. This is not entirely the case with other proposals such as in Chevallier [4] Bosma, Smeets [2] and the one hinted at in [7]. Unfortunately, we have not been able to carry out any experiments yet to see if the algorithm is practical in any sense.

We remark that the same idea and results would also work in finding small values of $|q + p_1\alpha_1 + \cdots + p_d\alpha_d|$. One would have to use the family of forms $p_1^2 + \cdots + p_d^2 + t(q + p_1\alpha_1 + \cdots + p_d\alpha_d)^2$ with $t \uparrow \infty$.

2. Quadratic forms

A quadratic form in the variables $x_1, x_2, \ldots, x_n$ is a homogeneous quadratic polynomial with coefficients in $\mathbb{R}$. We write such a form in the shape

$$\sum_{i,j=1}^{n} q_{ij} x_i x_j \quad \forall i, j : q_{ij} = q_{ji}.$$

Very often we abbreviate this to $Q(\mathbf{x})$. Without causing too much confusion we also use the notation Q for the $n \times n$-matrix with entries q_{ij}. We call this the matrix associated to the

quadratic form and the absolute value of the determinant of Q is called the *determinant* of the form. Notation: $D(Q)$.

The form $Q(\mathbf{x})$ is called *positive definite* if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and $Q(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$. From now on, when we speak of form, we mean a positive definite quadratic form.

For us, an important question will be to determine the minimal non-zero value of the set $\{Q(\mathbf{x})|\mathbf{x} \in \mathbb{Z}^n\}$, which we denote by $\mu(Q)$. We have the following theorem (see [3, Chapter 12]).

Theorem 2.1 (*Hermite*)**.** *For every $n \geq 2$ there exists γ_n such that $\mu(Q) \leq \gamma_n D(Q)^{1/n}$ for all positive definite forms Q in n variables.*

The smallest possible values of γ_n are known as *Hermite's constants*. We again denote them by γ_n. The first few values are $\gamma_2 = 2/\sqrt{3}$, $\gamma_3 = 2^{1/3}$, $\gamma_4 = \sqrt{2}, \ldots$. In general we have $\gamma_n \leq 2n/3$.

It is also interesting to consider the so-called successive minima of a form. The ith successive minimum $\mu_i(Q)$ is defined as the smallest real number ρ such that the ball defined by $Q(\mathbf{x}) \leq \rho$ contains a set of i independent vectors from $\mathbb{Z}^n$. In particular, $\mu_1(Q) = \mu(Q)$.

Another feature of forms in n variables is that the space of forms can be identified with the Riemannian symmetric space $O(n, \mathbb{R}) \setminus GL(n, \mathbb{R})$. The correspondence is given by the map $g \in GL(n, \mathbb{R}) \mapsto g^T g$, where g^T denotes the transpose of g. Notice that $g_1^T g_1 = g_2^T g_2$ if and only if there exists $u \in O(n, \mathbb{R})$ such that $g_2 = u g_1$. A metric on $O(n, \mathbb{R}) \setminus GL(n, \mathbb{R})$ is given by

$$ds^2 = \text{trace}((dY \cdot Y^{-1}) \cdot (dY \cdot Y^{-1})^T)$$

where $Y \in GL(n, \mathbb{R})$. The geodesics with respect to this metric are the one-dimensional families of quadratic forms

$$e^{\lambda_1 s} l_1(\mathbf{x})^2 + e^{\lambda_2 s} l_2(\mathbf{x})^2 + \cdots + e^{\lambda_n s} l_n(\mathbf{x})^2$$

parametrized by s, where $l_1, \ldots, l_n$ are independent forms. Thus we see that the family Q_t is a geodesic in the space of forms in $d + 1$ variables, which accounts for the name 'geodesic algorithm'.

3. Reduction of forms

Two forms $Q, \tilde{Q}$ in n variables are said to be equivalent if there exists $g \in GL(n, \mathbb{Z})$ such that $\tilde{Q}(\mathbf{x}) = Q(g\mathbf{x})$. It is common practice to choose suitably nice elements in each equivalence class, which we call reduced forms. There exist several notions of reduced forms, but for us the following two will be relevant: Minkowski reduced forms and LLL-reduced forms.

A form $Q(\mathbf{x})$ is called *Minkowski reduced* if for $i = 1, 2, 3, \ldots, n$ we have

$$Q(\mathbf{e}_i) \leq Q(\mathbf{m}) \quad \text{for all } \mathbf{m} \in \mathbb{Z} \text{ with } \gcd(m_i, \ldots, m_n) = 1. \tag{M}$$

Here $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is the standard basis of $\mathbb{R}^n$. Another way of stating these inequalities is to say that $Q(\mathbf{e}_i)$ is the minimum of all $Q(\mathbf{m})$ with $\mathbf{m} \in \mathbb{Z}^n$ such that $\mathbf{e}_1, \ldots, \mathbf{e}_{i-1}, \mathbf{m}$ can be extended to a basis of $\mathbb{Z}^n$. In particular $Q(\mathbf{e}_1) = q_{11}$ is the smallest non-zero value of Q restricted to $\mathbb{Z}^n$. The value $Q(\mathbf{e}_2) = q_{22}$ is the smallest value of all $\mathbf{x} \in \mathbb{Z}^n$ independent of $\mathbf{e}_1$. However, it is not always true that for a Minkowski reduced form $Q(\mathbf{e}_i)$ is the smallest value of Q at all $\mathbf{x} \in \mathbb{Z}^n$ independent of $\mathbf{e}_1, \ldots, \mathbf{e}_{i-1}$. The smallest value of i for which this fails is $i = 5$.

We denote the set of Minkowski reduced forms by $\mathcal{M}_n$. We have the following properties (see [3, Chapter 12]).

Proposition 3.1. *Let notations be as above. Then,*

1. *Every equivalence class of forms contains an element in* $\mathcal{M}_n$.
2. *Every equivalence class of forms contains finitely many forms in* $\mathcal{M}_n$.
3. *Two forms in the interior of* $\mathcal{M}_n$ *can only be equivalent via trivial substitutions of the form* $x_i \to \epsilon_i x_i$ *for all* i, *where* $\epsilon_i \in \{\pm 1\}$.
4. *For every* $Q \in \mathcal{M}_n$ *we have*
$$\mu_i(Q) \le Q(\mathbf{e}_i) \le 2^i \mu_i(Q), \quad i = 1, \ldots, n.$$
5. *The space* $\mathcal{M}_n$ *can be defined by a finite number of inequalities of the form* (M).

We illustrate the last fact for the cases $n = 2, 3$ (see [3, Chapter 12, Lemma 1.2]). When $n = 2$ the form reads $ax^2 + 2bxy + cy^2$ with associated matrix

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

The form is positive definite if and only if $a > 0, b^2 - ac < 0$. Its determinant reads $ac - b^2$. It is not hard to show that the form is Minkowski reduced if

$$|2b| \le a \le c.$$

A general ternary form (the case $n = 3$) reads

$$ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2.$$

One can show that it is Minkowski-reduced if and only if

$$a \le d \le f, \qquad |2b| \le a, \qquad |2c| \le a, \qquad |2e| \le d$$
$$a + d \ge (\pm b \pm c \pm e), \quad \text{zero or two minus signs.}$$

Having a Minkowski-reduced form equivalent to our given form Q yields a precious amount of information on Q. For example, the coefficient of x_1^2 of the reduced form is precisely $\mu(Q)$. Therefore it is of interest to find procedures that produce a reduced form equivalent to Q. In the case $n = 2$ there is already the well-known reduction procedure by Gauss. For other small values of n one can also device reduction procedures which are based on the inequalities that characterize Minkowski reducedness. Unfortunately it turns out that already when $n = 7$, the number of inequalities has risen to about 90,000. So it is clear that for $n > 6$ Minkowski reduction procedures tend to become unwieldy. Nevertheless, there are a number of papers in which one proposes Minkowski reduction algorithms for higher dimensions, see for example [1,6,13].

There is another concept of reducedness which avoids the exponential growth of reduction conditions, but at the cost of non-optimal output. It is called LLL-reduction, named after its inventors Laszlo Lovasz, Arjen Lentra and Hendrik Lenstra, who proposed it in 1982, [11]. The corresponding reduction algorithm that belongs to it has been extremely successful in many applications. It is simple, fast, even in large dimension, and yields good results.

To define LLL-reducedness we write Q in the form

$$\begin{aligned} Q(\mathbf{x}) = {} & b_1(x_1 + \mu_{12}x_2 + \cdots + \mu_{1n}x_n)^2 \\ & + b_2(x_2 + \mu_{23}x_3 + \cdots + \mu_{2n}x_n)^2 \\ & \vdots \\ & + b_{n-1}(x_{n-1} + \mu_{n-1,n}x_n)^2 + b_n x_n^2. \end{aligned}$$

We say that we have written Q in *recursive form*.

Definition 3.2. Fix a number $\omega \in [3/4, 1]$ (slack factor). We call the form Q *LLL-reduced* if

1. $|\mu_{ij}| \leq 1/2$ for all $i < j$.
2. $\omega b_i \leq b_{i+1} + \mu_{i,i+1}^2 b_i$ for all $i < n$. (Lovasz condition)

Using the recursive form Hermite already defines a notion of reduction, [8, p. 122 ff]. A form Q is called (Hermite) reduced if either one of the following holds,

- $n = 1$.
- When $n > 1$, we have $b_1 = \mu(Q)$, $|\mu_{1j}| \leq 1/2$ for $j = 2, \ldots, n$ and the form $Q - b_1 (x_1 + \mu_{12}x_2 + \cdots + \mu_{1n}x_n)^2$ in $x_2, \ldots, x_n$ is reduced.

Nowadays it is often called Hermite–Korkine–Zolotarev (HKZ) reducedness. One can easily show that HKZ-reducedness implies that $b_{i+1} + \mu_{i,i+1}^2 b_i \geq b_i$ for all $i < n$. So Lovasz condition can be seen as a relaxed version of this inequality (when $\omega < 1$).

In the literature LLL-reducedness is usually formulated in terms of lattice bases. In this paper we consider a version which is in terms of quadratic forms. Naively speaking one might think that reducedness of the quadratic form entails $|\mu_{ij}| \leq 1/2$ for all $i < j$, which we have seen earlier, and the condition $b_1 \leq b_2 \leq \cdots \leq b_n$. This is not going to work however. The innovation of LLL is to replace the naive condition $b_i \leq b_{i+1}$ by the Lovasz condition given above. Let us denote $\tilde{b}_i = b_{i+1} + \mu_{i,i+1}^2 b_i$. Then one can easily verify that if we swap the variables x_i, x_{i+1} in $Q(\mathbf{x})$ and rewrite the new form in recursive form again, the coefficient of $(x_i + \cdots)^2$ is precisely $\tilde{b}_i$. The coefficient of $(x_{i+1} + \cdots)^2$ in the new form is $b_{i+1}b_i/\tilde{b}_i$.

An LLL-reduced form has many interesting properties.

Theorem 3.3 (*LLL*)**.** *Let Q be a positive definite form in n variables and suppose Q is LLL-reduced with $\omega = 3/4$. Then*

1. $D(Q) \leq \prod_{i=1}^n Q(\mathbf{e}_i) \leq 2^{n(n-1)/2} D(Q)$.
2. $Q(\mathbf{e}_1) \leq 2^{(n-1)/2} D(Q)^{1/n}$.
3. $Q(\mathbf{e}_1) \leq 2^{n-1}\mu(Q)$.
4. *For $k = 1, 2, \ldots, n$ and all $j \leq k$ we have*
$$Q(\mathbf{e}_j) \leq 2^{n-1}\mu_k(Q).$$

Since the proofs in the literature are usually given for lattice bases we reproduce a proof valid for forms here.

Proof. First let us note that

$$Q(\mathbf{e}_i) = b_i + b_{i-1}\mu_{i-1,i}^2 + \cdots + b_1\mu_{1i}^2.$$

Secondly,

$$D(Q) = \prod_{i=1}^n b_i.$$

Thirdly, it follows from Lovasz condition that $b_i \geq (\omega - \mu_{i-1,i}^2) b_{i-1} \geq \frac{1}{2} b_{i-1}$.

By repeated application of the third inequality we find that $b_i \geq 2^{j-i} b_j$ whenever $j \leq i$, hence $b_j \leq 2^{i-j} b_i$. Together with our first observation this implies

$$\begin{aligned} Q(\mathbf{e}_i) &\leq b_i + \frac{1}{4}(b_{i-1} + \cdots + b_1) \\ &\leq b_i + \frac{1}{4}(2 + \cdots + 2^{i-1}) b_i \leq 2^{i-1} b_i. \end{aligned}$$

The first assertion of our theorem follows from

$$D(Q) = \prod_{i=1}^{n} b_i \le \prod_{i=1}^{n} Q(\mathbf{e}_i) \le \prod_{i=1}^{n} 2^{i-1} b_i = 2^{n(n-1)/2} D(Q).$$

To prove the second assertion we observe that for all $j \le i$,

$$Q(\mathbf{e}_j) \le 2^{j-1} b_j \le 2^{j-1} \cdot 2^{i-j} b_i = 2^{i-1} b_i.$$

Applied to the case $j = 1$ this gives

$$Q(\mathbf{e}_1)^n \le \prod_{i=1}^{n} 2^{i-1} b_i = 2^{n(n-1)/2} D(Q).$$

Taking the n-th roots gives our assertion.

The third assertion is a special case of the fourth, so we restrict to the fourth. Take a set of independent $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{Z}^n$ such that

$$\max(Q(\mathbf{x}_1), \ldots, Q(\mathbf{x}_k)) = \mu_k(Q).$$

Choose l minimal so that $\mathbf{x}_1, \ldots, \mathbf{x}_k$ lies in the span of $\mathbf{e}_1, \ldots, \mathbf{e}_l$. Since the $\mathbf{x}_i$ are independent we have $l \ge k$. Suppose $\mathbf{x}_i$ has l-th coordinate $\neq 0$. Denote this coordinate by ξ. Then ξ is a no-zero integer and we trivially get $Q(\mathbf{x}_i) \ge b_l \xi^2 \ge b_l$. So whenever $j \le l$,

$$Q(\mathbf{e}_j) \le 2^{l-1} b_l \le 2^{l-1} Q(\mathbf{x}_i) \le 2^{n-1} \mu_k(Q).$$

In particular, since $j \le l$, this assertion holds for all $j \le k$. □

For later purposes we also introduce partial LLL-reduction. We call the form Q *partially LLL-reduced* if

1. $|\mu_{i,i+1}| \le 1/2$ for all $i < n$.
2. $\omega b_i \le b_{i+1} + \mu_{i,i+1}^2 b_i$ for all $i < n$. (Lovasz condition)

We give a short description of the LLL-reduction algorithm for quadratic forms. There are two operations, a *shift* and a *swap*. A shift is a substitution of the form $x_r \to x_r + a x_s$ where $s > r$ and $a \in \mathbb{Z}$ is chosen such that the resulting μ_{rs} satisfies $|\mu_{rs}| \le 1/2$. A swap simply interchanges two neighbouring variables x_r, x_{r+1}.

Proposition 3.4. *Let Q be a form. We perform a shift or a swap and denote the resulting form by $\tilde{Q}$. Denote the parameters of the recursive form of $\tilde{Q}$ by $\tilde{b}_i$ and $\tilde{\mu}_{ij}$. Suppose we perform a shift $x_r \to x_r + a x_s$ (with $s > r$). Then*

1. $\tilde{b}_i = b_i$ *for all* i.
2. $\tilde{\mu}_{is} = \mu_{is} + a\mu_{ir}$ *for* $i = 1, \ldots, r-1$
3. $\tilde{\mu}_{rs} = \mu_{rs} + a$
4. $\tilde{\mu}_{ij} = \mu_{ij}$ *for all other* i, j.

Suppose we perform a swap $x_r \leftrightarrow x_{r+1}$. Then

1. $\tilde{b}_r = b_{r+1} + \mu_{r,r+1}^2 b_r$.
2. $\tilde{b}_{r+1} = b_r b_{r+1} / \tilde{b}_r$.
3. $\tilde{b}_i = b_i$ *for all* $i \neq r, r+1$.
4. $\tilde{\mu}_{ir} = \mu_{i,r+1}$ *for* $i < r$.
5. $\tilde{\mu}_{i,r+1} = \mu_{i,r}$ *for* $i < r$.

6. $\tilde{\mu}_{r,r+1} = b_r \mu_{r,r+1}/\tilde{b}_r$.
7. $\tilde{\mu}_{rj} = (b_r \mu_{r,r+1}\mu_{rj} + b_{r+1}\mu_{r+1,j})/\tilde{b}_r$ *for* $j > r+1$.
8. $\tilde{\mu}_{r+1,j} = \mu_{rj} - \mu_{r,r+1}\mu_{r+1,j}$ *for* $j > r+1$.
9. $\tilde{\mu}_{ij} = \mu_{ij}$ *for all other* i, j.

Proof. Straightforward computation. □

By a *global shift* we mean a sequence of shifts $x_i \to x_i + a x_j$ (with different a's) after which $|\mu_{ij}| \le 1/2$ for all $i < j$. Here is a possible implementation of LLL-reduction.

1. For $i = 1$ to $n-1$ perform a shift $x_i \to x_i + a x_{i+1}$. The result is that $|\mu_{i,i+1}| \le 1/2$ for $i = 1, \ldots, n-1$.
2. Enter the following *loop*: Find i such that $b_{i+1} + \mu_{i+1,i}^2 b_i < \omega b_i$.
   – If such i exists, swap x_{i+1} and x_i and perform the shifts $x_{i-1} \to x_{i-1} + a x_i$, $x_i \to x_i + a' x_{i+1}$ and $x_{i+1} \to x_{i+1} + a'' x_{i+2}$ (if they make sense). Then repeat the loop.
   – If such i does not exist: we exit the loop
3. The form is now partially LLL-reduced. To get an LLL-reduced form, perform a global shift.

The beauty of the LLL-algorithm is its running time.

Theorem 3.5 (*LLL*)**.** *Let* $B = \max_{i,j} |q_{ij}|$. *When* $\omega < 1$ *the number of loop-iterations of the algorithm is bounded above by* $n^2 \log(n^2 B/\mu(Q))/|\log \omega|$.

Of course it is a bit strange to have a running time estimate in terms of the unknown quantity $\mu(Q)$. However, in practice one works with integer quadratic forms, in which case we have $\mu(Q) \ge 1$.

We now give some explicit formula for q_i and μ_{ij} in terms of the coefficients q_{ij} of Q.

Theorem 3.6. *Let* Q *be a form in* n *variables with matrix* $(q_{ij})_{i,j=1,\ldots,n}$. *Let* b_i *and* μ_{ij} *be the coefficients corresponding to the descending shape of* Q. *For each* i, j *with* $1 \le i \le j \le n$ *we define*

$$B_{ij} = \begin{vmatrix} q_{11} & \cdots & q_{1,i-1} & q_{1j} \\ q_{21} & \cdots & q_{2,i-1} & q_{2j} \\ \vdots & & \vdots & \vdots \\ q_{i1} & \cdots & q_{i,i-1} & q_{ij} \end{vmatrix}.$$

Then $b_i = B_{i,i}/B_{i-1,i-1}$ *for* $i = 1, \ldots, n$ *where we adopt the convention* $B_{00} = 1$. *Moreover,* $\mu_{ij} = B_{ij}/B_{ii}$ *for all* i, j *with* $1 \le i < j \le n$.

Proof. We proceed by induction on i. For $i = 1$ the statement is straightforward to verify, all determinants have size 1×1. Now let $i > 1$ and suppose the statement holds for b_{i-1} and all $\mu_{i-1,j}$. Let us write

$$Q(x_1, \ldots, x_n) = b_1(x_1 + \mu_{12}x_2 + \cdots + \mu_{1n}x_n)^2 + \tilde{Q}(x_2, \ldots, x_n).$$

Note that the coefficients $\tilde{q}_{ij}$ of $\tilde{Q}$ are given by

$$\tilde{q}_{ij} = q_{ij} - q_{1i}q_{1j}/q_{11}$$

for all i, j with $2 \le i \le j \le n$. Denote by $\tilde{B}_{i,j}$ the determinant of the $(i-1) \times (i-1)$ matrix

$$\begin{pmatrix} \tilde{q}_{21} & \cdots & \tilde{q}_{2,i-1} & \tilde{q}_{1j} \\ \tilde{q}_{31} & \cdots & \tilde{q}_{3,i-1} & \tilde{q}_{2j} \\ \vdots & & \vdots & \\ \tilde{q}_{i1} & \cdots & \tilde{q}_{i,i-1} & \tilde{q}_{ij} \end{pmatrix}.$$

By induction we know that $q_i = \tilde{B}_{i,i}/\tilde{B}_{i-1,i-1}$ and $\mu_{ij} = \tilde{B}_{ij}/\tilde{B}_{ii}$ for $j > i$. Now consider the definition of B_{ij} given above. We perform a Gaussian row elimination using the element q_{11}. It is straightforward to see that we get $\tilde{B}_{ij} = B_{ij}/q_{11}$. This yields the desired formulae for q_i and μ_{ij}. □

Proposition 3.7. *Let notations be as in the previous theorem. Let $i < n$ and let C_i be the subdeterminant of $B_{i+1,i+1}$ obtained by deletion of the ith row and column. Then*

$$C_i B_{i,i} = B_{i+1,i+1} B_{i-1,i-1} + B_{i,i+1}^2.$$

As a consequence,

$$C_i/B_{i-1,i-1} = b_{i+1} + \mu_{i,i+1}^2 b_i.$$

Proof. The identity is an immediate consequence of the following general fact on determinant. Let M be an $n \times n$-matrix. Choose integers i, j such that $1 \le i < j \le n$. By $\tilde{M}$ we denote the $(n-2) \times (n-2)$-matrix obtained from M by deletion of the ith row and column and the jth row and column. By M_{kl} we denote the matrix obtained from M by deletion of the k-th row and l-th column. Then

$$\det(\tilde{M}) \det(M) = \det(M_{ii}) \det(M_{jj}) - \det(M_{ij}) \det(M_{ji}).$$

The proof of this identity is an interesting exercise in determinants. □

Corollary 3.8. *With the notations as above the LLL-reducedness conditions can be written as*

1. $2|B_{ij}| \le B_{ii}$ *for all* $1 \le i < j \le n$.
2. $\omega B_{i,i} \le C_i$ *for* $i = 1, \ldots, n-1$.

We can now rephrase Proposition 3.4 in terms of the determinants B_{ij}.

Proposition 3.9. *Let Q be a form in n variables and B_{ij} with $1 \le i \le j \le n$ its associated subdeterminants. After application of a substitution we denote the resulting form by $\tilde{Q}$ and its associated subdeterminants by $\tilde{B}_{ij}$.*

Suppose we apply a shift, that is we replace x_r by $x_r + a x_s$ for $a \in \mathbb{Z}$ and $s > r$. Then the subdeterminants associated to $\tilde{Q}$ read as follows,

1. $\tilde{B}_{is} = B_{is} + a B_{ir}$ *for* $i \le r$.
2. $\tilde{B}_{ij} = B_{ij}$ *for all other* i, j.
3. $\tilde{C}_r = C_r + 2a B_{rs} + a^2 B_{rr}$ *if* $r = s - 1$.
4. $\tilde{C}_i = C_i$ *whenever* $r \ne s-1$ *or* $r = s-1$ *and* $i \ne r$.

Suppose we apply the swap $x_r \leftrightarrow x_{r+1}$. Then the subdeterminants associated to $\tilde{Q}$ read as follows,

1. $\tilde{B}_{rr} = C_r$.
2. $\tilde{B}_{ir} = B_{i,r+1}$ *for all* $i < r$.

3. $\tilde{B}_{i,r+1} = B_{ir}$ *for all* $i < r$.
4. $\tilde{B}_{r,j} = (B_{r,r+1}B_{r,j} + B_{r-1,r-1}B_{r+1,j})/B_{rr}$ *for all* $j > r+1$.
5. $\tilde{B}_{r+1,j} = (B_{r+1,r+1}B_{r,j} - B_{r,r+1}B_{r+1,j})/B_{rr}$ *for all* $j > r+1$.
6. $\tilde{B}_{ij} = B_{ij}$ *for all other* i, j.
7. $\tilde{C}_r = B_{rr}$.
8. $\tilde{C}_{r-1} = (B_{r-2,r-2}C_r + B_{r-1,r+1}^2)/B_{r-1,r-1}$ *if* $r > 1$.
9. $\tilde{C}_{r+1} = (B_{r+2,r+2}C_r + \tilde{B}_{r+1,r+2}^2)/B_{r+1,r+1}$ *if* $r < n-1$.
10. $\tilde{C}_i = C_i$ *for all* $i \neq r-1, r, r+1$.

Proof. These are direct consequences of Propositions 3.4 and 3.7. □

We can now prove Theorem 3.5. During the LLL-algorithm we keep track of the product $\mathcal{B} = B_{11}B_{22}\cdots B_{nn}$. First we derive a lower bound for $\mathcal{B}$. Note that B_{ii} is the determinant of the form $Q(x_1, \ldots, x_i, 0, \ldots, 0)$ in i variables. Its smallest value is $\geq \mu(Q)$. So, by Theorem 2.1, we get $\mu(Q) \leq iB_{ii}^{1/i}$ (we used $\gamma_i \leq i$). Hence $B_{ii} \geq (\mu(Q)/i)^i$. Take the product over i to get $\mathcal{B} \geq (\mu(Q)/n)^{(n(n-1)/2)}$. An upper bound for B_{ii} can be given by $(Bi)^i$ (product of maximal lengths of columns). Hence $\mathcal{B} \leq (Bn)^{n(n-1)/2}$.

During the LLL-algorithm the value of $\mathcal{B}$ changes. From Proposition 3.9 it follows that the B_{ii} do not change after a shift. After a swap $x_r \leftrightarrow x_{r+1}$ all B_{ii} stay the same, except B_{rr} which becomes C_r. Since the swap is made we apparently have $C_r < \omega B_{rr}$, hence $\mathcal{B}$ is multiplied by a factor $<\omega$. Thus the maximal number of swaps can be bounded by

$$\frac{n(n-1)}{2}\frac{\log(Bn) - \log(\mu(Q)/n)}{|\log\omega|}$$

which yields the desired result.

4. Geodesic algorithms

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$. By Dirichlet's theorem there exist infinitely many $d+1$-tuples $p_1, \ldots, p_d, q \in \mathbb{Z}$ with $q > 0$ such that

$$\left|\alpha_i - \frac{p_i}{q}\right| \leq \frac{1}{q^{1+1/d}}, \quad i = 1, \ldots, d.$$

The goal of a continued fraction algorithm is to find such $d+1$-tuples or, if that is not possible, find approximations that come close to Dirichlet's inequalities. It is known that classical algorithms, such as the Jacobi–Perron algorithm, do not attain such quality of approximation. Recall for example the following theorem in the case $d = 2$.

Theorem 4.1 (*Schweiger*)**.** *There exists* $\delta > 0$ *such that for almost all pairs* α_1, α_2 *the Jacobi–Perron algorithm gives us*

$$\left|\alpha_i - \frac{p_i}{q}\right| \leq \frac{1}{q^{1+\delta}}, \quad i = 1, 2.$$

The optimal value of δ is not known, but experiments suggest that $\delta \approx 0.31$. In [10], Lagarias introduces another idea. Let $t > 0$ and consider the form

$$Q_t(\mathbf{x}, y) = (x_1 - \alpha_1 y)^2 + \cdots + (x_d - \alpha_d y)^2 + ty^2.$$

The important observation by Lagarias is the following.

Proposition 4.2. *Denote by* $|\mathbf{x}|$ *the Euclidean norm in* $\mathbb{R}^d$. *Suppose* $\mathbf{p} \in \mathbb{Z}^d$ *and* $q \in \mathbb{Z}_{\geq 0}$ *minimize the form* Q_t. *Then we have* $q > 0$ *and*

$$|q\boldsymbol{\alpha} - \mathbf{p}|q^{1/d} < \sqrt{d+1}.$$

Consequently, if $q > 0$,

$$\left|\alpha_i - \frac{p_i}{q}\right| \leq \frac{\sqrt{d+1}}{q^{1+1/d}}, \quad i = 1, \ldots, d.$$

Proof. The form Q_t has determinant t. So there exist $\mathbf{p} \in \mathbb{Z}^d$ and $q \in \mathbb{Z}_{\geq 0}$ such that

$$Q_t(\mathbf{p}, q) = |\mathbf{p} - \boldsymbol{\alpha} q|^2 + tq^2 \leq \gamma_{d+1} t^{1/(d+1)}.$$

Hence $|\mathbf{p} - \boldsymbol{\alpha} q|^2 \leq \gamma_{d+1} t^{1/(d+1)}$ and $q^{2/d} \leq \gamma_{d+1}^{1/d} t^{-1/(d+1)}$. Their product, and $\gamma_{d+1} < 2(d+1)/3$, yield the result. □

On this observation one can base the following algorithm to determine simultaneous rational approximations to $\alpha_1, \ldots, \alpha_d$ with the same denominator. Without loss of generality we can assume that $|\alpha_i| \leq 1/2$ for all i. We initialize with the form

$$Q_t^{(0)} = |\mathbf{x} - \boldsymbol{\alpha} y|^2 + ty^2$$

in the variables $x_1, \ldots, x_d$, y and $t \geq 1$. This form is Minkowski reduced and also LLL-reduced for any $\omega \leq 1$. We also define $P^{(0)}$ as the $(d+1) \times (d+1)$ identity matrix. We enter the following loop.

Loop:

- Determine the minimum of the set $\{t | Q_t^{(k)}$ is LLL-reduced$\}$ and call it t_k.
- Perform an LLL-reduction on $Q_{t_k - \epsilon}^{(k)}$ for infinitesimal $\epsilon > 0$ and let $\mathbf{x} \to A_k \mathbf{x}$ be the corresponding substitution of variables.
- Define $Q_t^{(k+1)}(\mathbf{x}) = Q_t^{(k)}(A_k \mathbf{x})$ and $P^{(k+1)} = P^{(k)} A_k$.

Remarks. 1. If we replace the word LLL-reduction with Minkowski reduction in the above algorithm we get the algorithm of Hermite and Lagarias.

2. For any k we have $Q_t(P^{(k)}\mathbf{x}) = Q_t^{(k)}(\mathbf{x})$. Set $(\mathbf{p}, q) = P^{(k)}\mathbf{e}_1$. Then, as a consequence of Theorem 3.3(2),

   $$|\mathbf{p} - q\boldsymbol{\alpha}|^2 + tq^2 \leq 2^{d/2} t^{1/(d+1)}.$$

   This implies that, if $q > 0$,

   $$|\mathbf{p} - q\boldsymbol{\alpha}| \leq 2^{d/4} q^{-1/d}.$$

   So the first column of $P^{(k)}$ gives us a simultaneous approximation to $\boldsymbol{\alpha}$ with a measure which differs from Dirichlet's approximation by at most a factor depending only on d.

3. We expect that most of the time the substitution-matrix A_k will simply be a shift or the result of a small number of LLL-steps.

In what follows we state a number of properties of the algorithm together with a number of theorems. Proofs will follow in the next section.

Verification of Minkowski reducedness involves the verification of a finite number of inequalities which are linear in the coefficients of the form. Although this is certainly not true for

LLL-reducedness, the *main observation of this paper* is that the inequalities to be verified for any $Q_t(M\mathbf{x})$ (with $M \in GL(d+1,\mathbb{Z})$) may not be linear in the coefficients of Q_t, but they turn out to be *linear in* t. Recall the LLL-conditions 3.8. They are stated in terms of the determinants B_{ij} and C_i formed out of the coefficients of $Q_t(M\mathbf{x})$, where $M \in GL(d+1,\mathbb{Z})$. Fortunately, these determinants have a form given by the following statement.

Proposition 4.3. *Consider the form $Q_t(\mathbf{x})$ defined above and let $M \in GL(d+1,\mathbb{Z})$. To $Q_t(M\mathbf{x})$ we associate the determinants B_{ij} as in Theorem 3.6 and C_i as in Proposition 3.7. Then each of these determinants is in $\mathbb{Z}[\alpha_1,\ldots,\alpha_d,t]$, they are quadratic in $\alpha_1,\ldots,\alpha_d$ and linear in t. Moreover, the coefficient of t is in $\mathbb{Z}$.*

Corollary 4.4. *Let notations be as above. Then the values of $t > 0$ for which $Q_t(M\mathbf{x})$ is LLL-reduced form an interval of the form $[t_0,t_1]\cap\mathbb{R}_{>0}$. More precisely, the set is either empty, or a point or a closed interval $t_0 \le t \le t_1$, or a half-open interval $0 < t \le t_1$.*

This is a direct consequence of the LLL-conditions 3.8 and Proposition 4.3. To determine the next value t_{k+1} in the algorithm we simply need to determine the largest $t < t_k$ such that at least one of the inequalities 3.8 becomes an equality. We then need to perform one or more shifts or swaps (or both). For each operation we need to update the determinants via the rules given in Proposition 3.9. Many of these rules are linear in the determinants but others are not. For example, consider the rule

$$\tilde{C}_{r-1} = (B_{r-2,r-2}C_r + B_{r-1,r+1}^2)/B_{r-1,r-1}.$$

Write

$$\tilde{C}_{r-1} = u_0t + v_0, \qquad C_r = u_1t + v_1$$
$$B_{r-2,r-2} = u_2t + v_2, \qquad B_{r-1,r+1} = u_3t + v_3, \qquad B_{rr} = u_4t + v_4$$

with $u_j \in \mathbb{Z}$ and $v_j \in \mathbb{Z}[\alpha_1,\ldots,\alpha_d]$. Then it follows from the equations that

$$u_0 = (u_1u_2 + u_3^2)/u_4, \qquad v_0 = (u_1v_2 + u_2v_1 + 2u_3v_3 - u_0v_4)/u_4.$$

So, although the update rules for the determinants are non-linear, the only non-linear part consists of division by an integer.

Here is a weak version of Theorem 2.1 in [10].

Theorem 4.5. *If $\boldsymbol{\alpha} \notin \mathbb{Q}^d$, the sequence of critical points $t_1, t_2, \ldots, t_k, \ldots$ is an infinite sequence descending to 0. If $\boldsymbol{\alpha} \in \mathbb{Q}^d$ the sequence $t_1, t_2, \ldots$ terminates at some value t_k.*

The following statement is actually Theorem 2.2 from [10], but with a different proof in the next section.

Theorem 4.6. *Suppose that the $\mathbb{Q}$-rank of the numbers $1,\alpha_1,\ldots,\alpha_d$ is r. Then for each i the limit $\lim_{t\downarrow 0}\mu_i(Q(\mathbf{x}))$ exists. Moreover, the limit is zero if $i \le r$ and it is positive if $i > r$.*

From this theorem and the properties of an LLL-reduced form in Theorem 3.3 it follows that the algorithm is capable of detecting linear relations between $1,\alpha_1,\ldots,\alpha_d$. On the other hand there are many properties that the LLL-based algorithm does not have in common with Lagarias' algorithm. For example, it does not guarantee that it finds Euclidean best approximations. Also, there is no analogue for Lagrange's theorem for ordinary continued fractions. Suppose we have an exceedingly good approximation in the sense that $\|q\boldsymbol{\alpha}\| \le \epsilon q^{-1/d}$ with very small ϵ. Setting

$t = \epsilon^2 q^{-2(d+1)/d}$ we see that this implies that

$$\mu(Q_t) \leq 2\epsilon^{2d/(d+1)} t^{1/(d+1)},$$

much smaller than the expected $\gamma_{d+1} t^{1/(d+1)}$. Suppose that during the algorithm the value t corresponds with the substitution matrix P, i.e. $Q_t(P\mathbf{x})$ is LLL-reduced. Then it follows from Theorem 3.3(2) that

$$Q_t(P\mathbf{e}_1) \leq 2^d \mu(Q_t) \leq 2^{d+1} \epsilon^{2d/(d+1)} t^{1/(d+1)}. \tag{S}$$

Unfortunately we cannot conclude from this that the vector $P\mathbf{e}_1$ corresponds to the excellent approximation $(\mathbf{p}, q)$ we are looking for. However, if

$$Q_t(P\mathbf{e}_2) > 2^d \max(2\epsilon^{2d/(d+1)} t^{1/(d+1)}, Q_t(P\mathbf{e}_1)),$$

we can conclude that $(\mathbf{p}, q) = P\mathbf{e}_1$. This is a consequence of Theorem 3.3(4) with $s = 2$. So under favourable circumstances we can determine excellent approximations.

It is well-known that the LLL-algorithm has been extremely successful in the explicit solution of diophantine equations (see [5]). The reason is that LLL is capable of showing the *non-existence* of excellent approximations in the sense that $\|q\boldsymbol{\alpha}\| \leq \epsilon q^{-1/d}$ with q less than a given Q. One simply has to verify that inequality (S) above is violated. However, for this one does not need the algorithm sketched above. A direct application of LLL will do.

If one is only interested in the vector $P^{(k)}\mathbf{e}_1$, one can skip a number of steps in the algorithm. From the update formulas one sees that the values of B_{ii}, $B_{i,i+1}$ and C_i are not affected if we perform a shift $x_r \to x_r + ax_s$ with $s > r + 1$. Nor are the LLL-conditions $2|B_{i,i+1}| \leq B_{ii}$ and $\omega B_{i,i} \leq C_i$ affected. Furthermore, the substitution matrix A corresponding to a shift has the property that $A\mathbf{e}_1 = \mathbf{e}_1$. These remarks suggest the following *partial continued fraction algorithm*. We initialize Q_{t_0} and $P^{(0)}$ as before. Then we enter the following loop.

Loop: Determine the minimum of the set $\{t | Q_t^{(k)}$ is partially LLL-reduced$\}$ and call it t_k^*. Perform a partial LLL-reduction on $Q_{t_k^*-\epsilon}^{(k)}$ for infinitesimal $\epsilon > 0$ and let $\mathbf{x} \to A_k\mathbf{x}$ be the corresponding substitution of variables. Define $Q_t^{(k+1)}(\mathbf{x}) = Q_t^{(k)}(A_k\mathbf{x})$ and $P^{(k+1)} = P^{(k)}A_k$.

The resulting sequence $t_1^*, t_2^*, \ldots$ is a subsequence of the sequence $t_1, t_2, \ldots$ we found earlier. If one wants, one can get an LLL-reduced version for any t by performing an additional global shift. However, if one is only interested in $P^{(k)}\mathbf{e}_1$ this is not necessary (shifts do not affect $\mathbf{e}_1$).

5. Proofs of statements

Proof of Proposition 4.3. The matrix corresponding to Q_t reads

$$Q_t = \begin{pmatrix} 1 & 0 & \cdots & -\alpha_1 \\ 0 & 1 & \cdots & -\alpha_2 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & -\alpha_d \\ -\alpha_1 & -\alpha_2 & \cdots & t + \alpha_1^2 + \cdots + \alpha_d^2 \end{pmatrix}.$$

We use the same notation Q_t for the matrix. Let us write $\tau = t + \alpha_1^2 + \cdots + \alpha_d^2$. The determinants B_{ij} and C_i are determinants of matrices which have the following form, $A^T Q_t B$ where A, B are $(d+1) \times k$ matrices of rank k and entries in $\mathbb{Z}$. There exists invertible $k \times k$-matrices R, S with

integer entries such that the last rows of AR and BS have at most one non-zero entry, which is at place k. The only entry in the matrix $R^T A^T Q_t B S$ which contains τ, is the one at place k, k. The entries at places i, j with $1 \le i, j \le k-1$ are integers. Hence the determinant is linear in τ and the coefficient of τ is an integer. This proves the second part of Proposition 4.3.

To prove the first part it suffices to show it after setting $\tau = 0$ (i.e. $t = -\alpha_1^2 - \cdots \alpha_d^2$). For this value of t the matrix Q_t is an integer matrix plus a rank 2 matrix with entries that are linear in $\alpha_1, \ldots, \alpha_d$. The same holds for the matrix $A^T Q_t B$. Hence its determinant is a quadratic polynomial in the α_i with integer coefficients. □

For the proof of Theorem 4.5 we need a lemma.

Lemma 5.1. *Let $t_0 > 0$. Then the number of $M \in GL(d+1, \mathbb{Z})$ such that $Q_t(M\mathbf{x})$ is LLL-reduced for some $t \ge t_0$, is finite.*

Proof. Note that the successive minima $\mu_i(Q_t)$ are decreasing in t. Let $t \in [t_0, 1]$ and suppose $M \in GL(d+1, \mathbb{Z})$ is such that $Q_t(M\mathbf{x})$ is LLL-reduced. Consider the following estimates for $i = 1, 2, \ldots, d+1$,

$$Q_{t_0}(M\mathbf{e}_i) \le Q_t(M\mathbf{e}_i) \le 2^d \mu_i(Q_t) \le 2^d \mu_i(Q_1).$$

The middle estimate follows from Theorem 3.3(4). Since the inequality $Q_t(\mathbf{x}) \le 2^d \mu_i(Q_1)$ in $\mathbf{x} \in \mathbb{Z}^{d+1}$ has finitely many solutions, there are finitely many possibilities for $M\mathbf{e}_i$, the ith column of M. Hence our lemma follows. □

Proof of Theorem 4.5. Recall that the t_i form a decreasing sequence. Suppose there is a point of accumulation $t_\infty > 0$. Hence we have infinitely many $t_k > t_\infty$ such that $Q_{t_k}(P^{(k)}\mathbf{x})$ is LLL-reduced. Since the $P^{(k)}$ are distinct, this contradicts the lemma we just proved. So we have either $\lim_{k\to\infty} t_k = 0$ or the sequence $t_1, t_2, \ldots$ stops at t_k. In the latter case the form $Q_t(P^{(k)}\mathbf{x})$ is LLL-reduced for all t with $0 < t < t_k$. Since $Q_t(P^{(k)}\mathbf{e}_1) \le Ct^{1/(d+1)}$ for some $C > 0$ we see that $Q_0(P^{(k)}\mathbf{e}_1) = 0$. Letting $(p_1, \ldots, p_d, q) = P^{(k)}\mathbf{e}_1$ this implies that $(p_1 - \alpha_1 q)^2 + \cdots + (p_d - \alpha_d q)^2 = 0$. Hence $p_i/q = \alpha_i$ for $i = 1, \ldots, d$. So all α_i are rational.

Similarly, if the sequence of t_i is infinite, we let $(\mathbf{p}^{(k)}, q) = P^{(k)}\mathbf{e}_1$ and see that $|\mathbf{p}^{(k)} - \boldsymbol{\alpha} q| \to 0$ as $k \to \infty$. This is only possible if not all α_i are rational. □

Proof of Theorem 4.6. It suffices to prove that $\lim_{t\downarrow 0} \mu_{r+1}(Q_t) > 0$ if $r \le d$ and $\lim_{t\downarrow 0} \mu_r(Q_t) = 0$.

Let $L \subset \mathbb{Z}^{d+1}$ be the lattice of vectors $(l_0, \mathbf{l})$ such that $l_0 + \mathbf{l} \cdot \boldsymbol{\alpha} = 0$. It has $\mathbb{Z}$-rank $d + 1 - r$. We choose a fixed basis B. Suppose we have $r+1$ independent vectors $(\mathbf{p}_i, q_i) \in \mathbb{Z}^{d+1}$ such that $Q_t(\mathbf{p}_i, q_i) \le \mu_{r+1}(Q_t)$ for $i = 1, \ldots, r+1$. Then there exists a vector $(l_0, \mathbf{l}) \in B$ and an i such that $l_0 q_i + \mathbf{l} \cdot \mathbf{p}_i \ne 0$. Hence

$$1 \le |l_0 q_i + \mathbf{l} \cdot \mathbf{p}_i| = |\mathbf{l} \cdot (\mathbf{p}_i - q_i \boldsymbol{\alpha})|.$$

This implies $|\mathbf{p}_i - q_i \boldsymbol{\alpha}| \ge 1/|\mathbf{l}|$ and hence

$$\mu_{r+1}(Q_t) \ge Q_t(\mathbf{p}_i, q_i) \ge 1/|\mathbf{l}|^2.$$

This proves the first part of Theorem 4.6.

We may assume without loss of generality that $1, \alpha_1, \ldots, \alpha_{r-1}$ are $\mathbb{Q}$-linear independent. For any $j \ge 1$ we write $\alpha_j = a_{j0} + \sum_{k=1}^{r-1} a_{jk}\alpha_k$ and let N be the common denominator of the a_{jk}.

Claim. *To any $\epsilon > 0$ there exists a rank r matrix $P = (p_{ij})_{i=1,\ldots,r;\, j=0,\ldots,r-1}$ such that $|p_{ij} - \alpha_j p_{i0}| \le \epsilon$ for every $i = 1, \ldots, r$; $j = 1, \ldots, r-1$.*

We then proceed as follows. Define $p_{ij} = a_{j0}p_{i0} + \sum_{k=1}^{r-1} a_{jk}p_{ik}$ for any $j \geq 1$ and $i = 1, \ldots, r$. Then

$$|Np_{ij} - N\alpha_j p_{i0}| = \left|\sum_{k=1}^{r-1} Na_{jk}(p_{ik} - \alpha_k p_{i0})\right| \leq rNA\epsilon$$

where $A = \max_{i,j} |a_{jk}|$. Write $p_{i0} = q_i$ and $\mathbf{p}_i = (p_{i1}, \ldots, p_{id})$ for $i = 1, \ldots, r$. Note that $N\mathbf{p}_i \in \mathbb{Z}^d$ and $Nq_i \in \mathbb{Z}$. Then,

$$\mu_r(Q_t) \leq \max_i(|N\mathbf{p}_i - q_i\boldsymbol{\alpha}|^2 + tq_i^2) \leq d(rNA\epsilon)^2 + t\max_i q_i^2.$$

Letting t go to 0 this implies $\lim_{t\downarrow 0} \mu_r(Q_t) \leq d(rNA\epsilon)^2$. Since ϵ can be chosen arbitrarily small our second assertion follows. □

Acknowledgements

Many thanks to Robbert Fokkink and Cor Kraaikamp for their invitation to the workshop 'Probability and Numbers' in Delft. Thanks also to Catherine Goldstein who provided me with a number of very interesting references to the work of Charles Hermite.

References

[1] L. Afflerbach, H. Grothe, Calculation of Minkowski-reduced lattice bases, Computing 35 (1985) 269–276.
[2] W. Bosma, I. Smeets, Finding simultaneous diophantine approximations with prescribed quality, in: ANTS X: Proceedings of the Tenth Algorithmic Number Theory Symposium, UC San Diego 2012, in: The Open Book Series, vol. 1, 2013, See also http://arxiv.org/pdf/1001.4455v1.pdf.
[3] J.W.S. Cassels, Rational Quadratic Forms, Academic Press, 1968.
[4] N. Chevallier, Best simultaneous diophantine approximation and multidimensional continued fraction algorithms, Mosc. J. Combin. Number Theory 3 (2013) 3–36.
[5] B.M.M. de Weger, Solving exponential diophantine equations using lattice basis reduction algorithms, J.Number Theory 26 (1987) 325–367.
[6] B. Helfrich, Algorithms to construct Minkowski reduced and Hermite reduced lattice bases, Theoret. Comput. Sci. 41 (1985) 125–139.
[7] D. Hensley, Continued Fractions, World Scientific, 2006.
[8] Ch. Hermite, Extraits de lettres de M. Hermite à M. Jacobi sur différents objets de la théorie des nombres, J. Reine Angew. Math. 40 (1850) 261–315;
E. Picard (Ed.), Repr. (with corrections) in Hermite, OEuvres, Vol. 1, Gauthier-Villars, Paris, 1905, pp. 100–163.
[9] G. Humbert, Sur la méthode d'approximation d'Hermite, J. Math. Pures Appl. 2 (1916) 79–103.
[10] J.C. Lagarias, Geodesic multidimensional continued fractions, Proc. Lond. Math. Soc. 69 (1994) 464–488.
[11] A.J. Lenstra, H.W. Lenstra, L. Lovasz, Factoring polynomials with rational coefficients, Math. Ann. 261 (1982) 515–534.
[12] E. Picard, Leçon sur l'Œuvre scientifique de Charles Hermite. Repr. préface, in: Hermite's Collected Works (1905–1917), Vol. I, 1901, p. vii-xl.
[13] Wen Zhang, San-zheng Qiao, Yi-min Wei, HKZ and Minkowski reduction algorithms for lattice-reduction aided MIMO detection, IEEE Trans. Signal Process. 60 (2012) 5963–5976.