

# An Optimal, Stable Continued Fraction Algorithm for Arbitrary Dimension

Carsten Rössner and Claus P. Schnorr

Fachbereich Mathematik/Informatik  
Universität Frankfurt  
PSF 11 19 32  
D-60054 Frankfurt am Main, Germany  
{roessner,schnorr}@cs.uni-frankfurt.de

**Abstract.** We analyse a continued fraction algorithm (abbreviated CFA) for arbitrary dimension  $n$  showing that it produces simultaneous diophantine approximations which are up to the factor  $2^{(n+2)/4}$  best possible. Given a real vector  $x = (x_1, \dots, x_{n-1}, 1) \in \mathbb{R}^n$  this CFA generates a sequence of vectors  $(p_1^{(k)}, \dots, p_{n-1}^{(k)}, q^{(k)}) \in \mathbb{Z}^n$ ,  $k = 1, 2, \dots$  with increasing integers  $|q^{(k)}|$  satisfying for  $i = 1, \dots, n-1$

$$|x_i - p_i^{(k)}/q^{(k)}| \leq 2^{(n+2)/4} \sqrt{1 + x_i^2} / |q^{(k)}|^{1 + \frac{1}{n-1}}.$$

By a theorem of Dirichlet this bound is best possible in that the exponent  $1 + \frac{1}{n-1}$  can in general not be increased.

## 1 Introduction

We analyse a CFA which computes for real vectors  $x \in \mathbb{R}^n$  diophantine approximations to  $x$  that are up to the factor  $2^{(n+2)/4}$  best possible. Given  $x \in \mathbb{R}^n$  this CFA constructs a sequence of lattice bases of the lattice  $\mathbb{Z}^n$  consisting of vectors that approximate the line  $x\mathbb{R}$ . For given  $\epsilon > 0$ , this CFA either finds an integer relation  $m \in \mathbb{Z}^n - 0$  for  $x$ , i.e.  $\langle m, x \rangle = 0$ , of Euclidean length at most  $2^{n/2} \epsilon^{-1}$  or it proves that no integer relation of length  $\leq \epsilon^{-1}$  exists. For this the algorithm uses  $O(n^4 (n + |\log \epsilon|))$  arithmetic operations on real numbers with exact arithmetic. For a rational input vector  $x := (q_1, \dots, q_n)/q_n$ , with  $q_1, \dots, q_n \in \mathbb{Z}$  the algorithm has polynomial bit complexity in the input size  $\sum_{i=1}^n [\log |q_i|] + |\log \epsilon|$ . Our analysis relies on the *dual* lattice basis which we show to consist of very short vectors, see Theorems 1, 2. From this we greatly improve the known bounds for the primary lattice basis and for diophantine approximation. The crucial role of the dual basis escaped in all previous studies.

Our algorithm is a variant of the HJLS-algorithm of Hastad, Just, Lagarias, Schnorr [HJLS89] for finding integer relations for a real vector  $x$  which in turn relies on the algorithms of Bergman [Berg80], Ferguson, Forcade [FF79] and Lenstra, Lenstra, Lovász [LLL82]. It also incorporates ideas of Just [Ju92], Ferguson and Bailey [FB92] and Rössner, Schnorr [RS95]. We present a stable floating point version of this algorithm, prove stability in Theorem 6 and demonstrate its stability by experimental data.

The problem of higher dimensional CFA has been widely studied by Jacobi [Ja1868], Perron [Pe1907], Bernstein [Bern71], Szekeres [Sz70], Ferguson, Forcade [FF79], Bergman [Berg80] and Lenstra, Lenstra, Lovász [LLL82]. The HJLS–algorithm of [HJLS89] is a variant of the algorithms in [FF79], [Berg80] and [LLL82]. It finds short integer relations for  $x$  in polynomial time using exact arithmetic on real numbers. Just [Ju92] showed that a variant of this algorithm provides diophantine approximations satisfying  $|x_i - p_i^{(k)}/q^{(k)}| \leq 2^{(n+2)/4} \sqrt{1 + x_i^2}/|q^{(k)}|^{1 + \frac{1}{2n(n-1)}}$ . We improve the analysis of [Ju92].

Ferguson and Bailey [FB92] have implemented a close variant of the HJLS–algorithm which they call the PSLQ–algorithm. Their experimental results show that this CFA produces simultaneous diophantine approximations that are far better than for any other known algorithm. Recently Bailey, Borwein, Plouffe [BBP96] found surprising new approximation algorithms for  $\pi$ ,  $\ln(2)$  using this CFA. While this CFA could so far not be analyzed we prove for the first time the superiority of this CFA.

## 2 Preliminaries

Let  $\mathbb{R}^n$  be the  $n$ –dimensional real vector space equipped with the ordinary inner product  $\langle \cdot, \cdot \rangle$  and Euclidean length  $\|y\| := \langle y, y \rangle^{1/2}$ . We let  $[y_1, \dots, y_m]$  denote the matrix with column vectors  $y_1, \dots, y_m$  and  $[\cdot]$  is the nearest integer function to a real number  $r$ ,  $[r] = [r + 0.5]$ .

A non–zero vector  $m \in \mathbb{Z}^n$  is called an *integer relation* for  $x \in \mathbb{R}^n$  if  $\langle x, m \rangle = 0$ . We let  $\lambda(x)$  denote the length  $\|m\|$  of the shortest integer relation  $m$  for  $x$ ,  $\lambda(x) = \infty$  if no relation exists.

Throughout this paper,  $b_1, \dots, b_n$  is an ordered basis of the integer lattice  $\mathbb{Z}^n$  and its *dual* basis  $a_1, \dots, a_n$  is defined by  $[a_1, \dots, a_n]^\top := [b_1, \dots, b_n]^{-1}$ . Let  $x \in \mathbb{R}^n$  be a non–zero vector, set  $b_0 := x$ . We associate with the basis  $b_1, \dots, b_n$  the orthogonal projections

$$\begin{aligned} \pi_{i,x} : \mathbb{R}^n &\longrightarrow \text{span}(x, b_1, \dots, b_{i-1})^\perp && \text{and} \\ \pi_i : \mathbb{R}^n &\longrightarrow \text{span}(b_1, \dots, b_{i-1})^\perp && \text{for } i = 1, \dots, n, \end{aligned}$$

where  $\text{span}(b_j, \dots, b_{i-1})$  denotes the linear space generated by  $b_j, \dots, b_{i-1}$  and  $\text{span}(b_j, \dots, b_{i-1})^\perp$  its orthogonal complement in  $\mathbb{R}^n$ . We abbreviate  $\widehat{b}_{i,x} := \pi_{i,x}(b_i)$  and  $\widehat{b}_i := \pi_i(b_i)$ . The vectors  $\widehat{b}_{1,x}, \dots, \widehat{b}_{n,x}$  (resp.  $\widehat{b}_1, \dots, \widehat{b}_n$ ) are pairwise orthogonal. They are called the *Gram–Schmidt orthogonalization* of  $x, b_1, \dots, b_n$  (resp.  $b_1, \dots, b_n$ ). The *Gram–Schmidt coefficients*  $\mu_{i,j}$  of the factorization  $[x, b_1, \dots, b_n] = [x, \widehat{b}_{1,x}, \dots, \widehat{b}_{n,x}] (\mu_{i,j})_{0 \leq i, j \leq n}^\top$  are defined as  $\mu_{i,j} := \langle b_i, \widehat{b}_{j,x} \rangle / \|\widehat{b}_{j,x}\|^2$ . If  $\widehat{b}_{j,x} = 0$ , we set  $\mu_{i,j} = 0$  for  $i \neq j$  and  $\mu_{j,j} = 1$ . The matrix  $(\mu_{i,j})_{0 \leq i, j \leq n}$  is lower triangular with all diagonal elements 1. Finally we note that  $a_n = \widehat{b}_n / \|\widehat{b}_n\|^2$  since both  $a_n$  and  $\widehat{b}_n$  are orthogonal to  $b_1, \dots, b_{n-1}$ .

The (ordered) vectors  $x, b_1, \dots, b_n$  are *size–reduced* if  $|\mu_{k,j}| \leq \frac{1}{2}$  holds for  $1 \leq j < k \leq n$  and  *$L^3$ –reduced* if they are size–reduced and the inequality

$\frac{3}{4} \|\pi_{k-1,x}(b_{k-1})\|^2 \leq \|\pi_{k-1,x}(b_k)\|^2$  holds for  $k = 2, \dots, n$ . If  $L^3$ -reduced the vectors satisfy  $\|\widehat{b}_{i,x}\|^2 \leq 2 \|\widehat{b}_{i+1,x}\|^2$  for  $i = 1, \dots, n-1$ .

**Models of Computation.** We distinguish three models of computation for the CFA.

*Exact Real Arithmetic.* For real input  $x \in \mathbb{R}^n$  we use exact arithmetic over real numbers. This version of the CFA can use either Gram–Schmidt orthogonalization or Givens Rotation with square roots. The analysis of the HJLS–algorithm applies.

*Exact Integer Arithmetic.* For rational input  $x \in \mathbb{Q}^n$  we can use exact arithmetic over the integers. The rational numbers  $\mu_{i,j}, \|\widehat{b}_{j,x}\|^2$  are represented by their numerator and denominator. This version of the CFA uses Gram–Schmidt orthogonalization. The analysis of the  $L^3$ -algorithm [LLL82] for lattice basis reduction applies.

*Floating Point Arithmetic.* For rational input  $x$  we can speed up the CFA in that we replace the exact arithmetic on the rationals  $\mu_{i,j}, \|\widehat{b}_{j,x}\|$  by floating point arithmetic. The vectors  $x, b_1, \dots, b_n, a_1, \dots, a_n$  are kept in exact representation. In order to minimize floating point errors we use, instead of the  $\mu_{i,j}$ , the normalized coefficients  $\tau_{i,j} := \mu_{i,j} \|\widehat{b}_{j,x}\|$ . We call the entities  $\tau_{i,j}$  for  $0 \leq i, j \leq n$  the *orthonormalization* of  $x, b_1, \dots, b_n$ . Note that  $\tau_{i,i} = \|\widehat{b}_{i,x}\|$ . The  $L^3$ -property  $\frac{3}{4} \|\pi_{k-1,x}(b_{k-1})\|^2 \leq \|\pi_{k-1,x}(b_k)\|^2$  is expressed by  $\frac{3}{4} \tau_{k-1,k-1}^2 \leq \tau_{k,k}^2 + \tau_{k,k-1}^2$ . The  $\tau_{i,j}$  are not rational but require square roots, we compute them in floating point arithmetic using Givens Rotation.

We present our algorithm in its floating point version. From this description the details for the other models of computation are straightforward and left to the reader.

### 3 The Algorithm Description

This algorithm improves the HJLS–algorithm [HJLS89] towards numerical stability. Given a real vector  $x \in \mathbb{R}^n$  and  $\epsilon > 0$ , the HJLS–algorithm either finds an integer relation  $m$  for  $x$  with  $\|m\| \leq 2^{n/2-1} \min\{\lambda(x), \epsilon^{-1}\}$  or it proves  $\lambda(x) \geq \epsilon^{-1}$ . The HJLS–algorithm performs reduction and exchange steps on the linearly dependent system of vectors of the matrix  $x, b_1, \dots, b_n$  where initially  $b_1, \dots, b_n$  are set to the unit vectors in  $\mathbb{R}^n$ . The vector  $x$  remains unchanged and the vectors  $b_1, \dots, b_n$  remain a basis of the lattice  $\mathbb{Z}^n$ . The HJLS–algorithm uses exact arithmetic on real numbers. Its reduction and exchange steps minimize  $\max_{1 \leq i \leq n} \|\widehat{b}_{i,x}\|$ .

The HJLS–algorithm terminates if either  $x \in \text{span}(b_1, \dots, b_{n-1})$ , i.e. if a swap  $b_{n-1} \leftrightarrow b_n$  results in  $\widehat{b}_{n-1,x} = 0$ , or if  $\max_{1 \leq i \leq n} \|\widehat{b}_{i,x}\| \leq \epsilon$ . In the first case, the last vector  $a_n$  of the dual basis is an integer relation for  $x$ . In the latter case, we have  $\lambda(x) \geq \epsilon^{-1}$  which follows from

[HJLS89] **Proposition 3.1.** *Every basis  $b_1, \dots, b_n$  of  $\mathbb{Z}^n$  satisfies*

$$\lambda(x) \geq 1 / \max_{1 \leq i \leq n} \|\widehat{b}_{i,x}\|. \quad (1)$$

Our main modifications of the HJLS–algorithm are as follows:

1. We iteratively swap vectors  $b_{k-1}$ ,  $b_k$  with  $2 \leq k \leq n-1$  that do not satisfy the  $L^3$ –condition  $\frac{3}{4} \|\pi_{k-1,x}(b_{k-1})\|^2 \leq \|\pi_{k-1,x}(b_k)\|^2$ . The selection of  $k$ , either minimal as in the  $L^3$ –algorithm or so that  $i := k$  maximizes  $2^i \|\widehat{b}_{i,x}\|^2$  as proposed by Bergman, is irrelevant.

2. Before swapping the last two vectors  $b_{n-1}$  and  $b_n$  the basis  $\pi_{1,x}(b_1), \dots, \pi_{1,x}(b_n)$  is  $L^3$ –reduced. Here we follow Just [Ju92]. In the model of exact real arithmetic our algorithm essentially coincides with the algorithm of Just [Ju92] and her analysis of diophantine approximation properties applies.

3. We apply reduction in size, i.e. we reduce  $b_k$  so that  $|\mu_{k,i}| \leq 1/2$  for  $i = 1, \dots, k-1$ . Reduction in size has been neglected in [HJLS89] since it is pointless for the exact real arithmetic.

4. In the floating point version orthonormalization of the vectors  $x, b_1, \dots, b_n$  is done by Givens Rotation with a floating point error that is linear in  $n$  and  $\max_{0 \leq i \leq n} \|b_i\|$ , see [Ge75, GL89]. Givens Rotations has already been used in the parallel  $L^3$ –algorithms of Heckler, Thiele [HT93] and Joux [Jo93]. Ferguson and Bailey [FB92] essentially use Givens Rotation in connection with the HJLS–algorithm.

The test on  $\tau_{n,n} \neq 0$  actually checks whether  $\tau_{n,n} > 2^{-r}$  where  $r$  is the number of precision bits of the floating point arithmetic.

### Stable Continued Fraction Algorithm (SCFA)

INPUT  $x \in \mathbb{R}^n - 0$ ,  $\epsilon > 0$ .

1. *Initiation.* Let  $b_i \in \mathbb{Z}^n$  be the  $i$ –th unit vector. Compute the orthonormalization  $\tau_{i,j}$  for  $0 \leq i, j \leq n$  of  $x, b_1, \dots, b_n$  using Givens Rotation (see Section 5).  $s := 1$ .

2.  $L^3$ –reduction of  $\pi_{1,x}(b_1), \dots, \pi_{1,x}(b_{n-1})$ .

While there exists  $k$  with  $1 < k < n$  and  $\frac{3}{4} \tau_{k-1,k-1}^2 > \tau_{k,k}^2 + \tau_{k,k-1}^2$  size–reduce  $b_k$  with respect to  $b_{k-1}$  by setting  $b_k := b_k - \lceil \tau_{k,k-1} / \tau_{k-1,k-1} \rceil b_{k-1}$ , swap  $b_{k-1}$  and  $b_k$  and update the orthonormalization using Givens Rotation.

Reduce  $b_1, \dots, b_n$  in size. While  $|\tau_{s,s}| \leq \epsilon$  increment  $s$  to  $s+1$ .

Output  $(p_1, \dots, p_{n-1}, q) := b_1$ , the next approximation for  $x$ , see Theorem 3.

3. *Swap the last vectors.* Swap  $b_{n-1}$  and  $b_n$ , and update the orthonormalization using Givens Rotation. If  $\tau_{n,n} = 0$  and  $s < n$  then goto 2.

4. *Termination.* Compute  $[a_1, \dots, a_n]^\top := [b_1, \dots, b_n]^{-1}$ .

If  $\tau_{n,n} > 0$  a relation for  $x$  is found. Output the nearby point  $x' := x$  and the relation  $a_n$  for  $x$ .

If  $s = n$  then  $\tau_{i,i} \leq \epsilon$  holds for  $i = 1, \dots, n$ .

Compute  $\pi_n(x) \in \text{span}(b_1, \dots, b_{n-1})^\perp$ , output the nearby point  $x' := x - \pi_n(x)$ , the relation  $a_n$  for  $x'$  and “ $\lambda(x) \geq \epsilon^{-1}$ ”.

If  $\epsilon = 0$  then SCFA produces a possibly infinite sequence of vectors  $b_1$ , occurring after  $L^3$ –reduction, that are good diophantine approximations to  $x$ .

**Correctness Properties.** 1. Upon termination of step 2 we have  $\tau_{i,i} \leq \epsilon$  for  $i = 1, \dots, s-1$  and the projected basis  $\pi_{1,x}(b_1), \dots, \pi_{1,x}(b_{n-1})$  is  $L^3$ -reduced. 2. Before swapping  $b_{n-1}$  and  $b_n$  we have  $s < n$  (note that  $s \neq n$  since  $\tau_{n,n} = 0$ ) and  $\tau_{s,s}^{-1} < \epsilon^{-1}$ . Therefore the  $L^3$ -reducedness of  $\pi_{1,x}(b_1), \dots, \pi_{1,x}(b_{n-1})$  implies that  $\|\widehat{b}_{n-1,x}\|^{-1} < 2^{(n-1-s)/2} \epsilon^{-1}$ . 3. We recall from [RS95], Theorem 6 that SCFA computes a nearby point  $x'$  and a non-zero vector  $a_n \in \mathbb{Z}^n$  so that  $\langle a_n, x' \rangle = 0$  and  $\lambda(\bar{x}) \geq \epsilon^{-1}/2$  holds for all  $\bar{x} \in \mathbb{R}^n$  with  $\|x - \bar{x}\| < \|x - x'\|/2$ . If  $x' \neq x$  we have  $\lambda(x) \geq \epsilon^{-1}$ .

## 4 Analysis of SCFA in Exact Real Arithmetic

**Theorem 1.** *Throughout the computation we have  $\|a_n\| \leq 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\}$ . Moreover,  $\|a_n\| \leq 2^{n/2+1} \lambda(x')$  holds upon termination.*

For the first time we prove in Theorem 1 that SCFA outputs a relation  $a_n$  for the nearby point  $x' \neq x$  which has, up to the factor  $2^{n/2+1}$ , minimal length.

**Proof.** We let  $\bar{b}_1, \dots, \bar{b}_n, \bar{a}_1, \dots, \bar{a}_n$  denote the dual bases before and  $b_1, \dots, b_n, a_1, \dots, a_n$  after an arbitrary swap  $b_{n-1} \leftrightarrow b_n$  of SCFA. Let  $\bar{\mu}_{i,j}$  be the Gram-Schmidt coefficients and  $\widehat{b}_{i,x}$  be the orthogonal vectors of  $x, \bar{b}_1, \dots, \bar{b}_n$ . We have  $\widehat{b}_{n-1,x} = \bar{\mu}_{n,n-1} \widehat{b}_{n-1,x}$  with  $|\bar{\mu}_{n,n-1}| \leq \frac{1}{2}$ .

From the characterization of  $a_{n-1}$  as  $\langle a_{n-1}, b_i \rangle = \delta_{n-1,i}$ , which holds throughout the algorithm, we infer that

$$a_{n-1} = \frac{\widehat{b}_{n-1,x}}{\|\widehat{b}_{n-1,x}\|^2} - \frac{\langle b_n, \widehat{b}_{n-1,x} \rangle}{\|\widehat{b}_{n-1,x}\|^2} a_n.$$

Applying this equation to the vectors  $\bar{b}_1, \dots, \bar{b}_n, \bar{a}_1, \dots, \bar{a}_n$  and  $|\bar{\mu}_{n,n-1}| \leq \frac{1}{2}$  implies the recursion formula

$$\begin{aligned} \|a_n\| &= \|\bar{a}_{n-1}\| = \|\widehat{b}_{n-1,x}\|^{-1} + |\bar{\mu}_{n,n-1}| \|\bar{a}_n\| \\ &\leq \|\widehat{b}_{n-1,x}\|^{-1} + \frac{1}{2} \|\bar{a}_n\|. \end{aligned}$$

From the  $L^3$ -reducedness of  $\pi_{1,x}(\bar{b}_1), \dots, \pi_{1,x}(\bar{b}_{n-1})$  and inequality (1) we see that

$$2^{(n-1)/2} \|\widehat{b}_{n-1,x}\| \geq \max_{1 \leq i \leq n} 2^{i/2} \|\widehat{b}_{i,x}\| \geq 2^{1/2} \lambda(x)^{-1}.$$

Using  $\|\widehat{b}_{n-1,x}\|^{-1} \leq 2^{n/2-1} \lambda(x)$  and  $\|\widehat{b}_{n-1,x}\|^{-1} \leq 2^{n/2-1} \epsilon^{-1}$ , which follows from correctness property 2, we can rewrite the recursion formula to

$$\|a_n\| \leq 2^{n/2-1} \min\{\epsilon^{-1}, \lambda(x)\} + \frac{1}{2} \|\bar{a}_n\|.$$

This inequality holds for every exchange  $b_{n-1} \leftrightarrow b_n$  of SCFA. Suppose that there are exactly  $t$  such exchanges and using that initially  $\|a_n\| = 1$  we obtain the first claim:

$$\|a_n\| \leq 2^{n/2-1} \min\{\epsilon^{-1}, \lambda(x)\} \sum_{j=0}^{t-1} 2^{-j} + 2^{-t} \leq 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\} .(2)$$

Since the inequality  $\|a_n\| \leq 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\}$  holds after any swap  $b_{n-1} \leftrightarrow b_n$  it must always hold because  $a_n$  does not change between two swaps.

It remains to prove the second claim in the case  $x' \neq x$ . For this we use

**[RS95] Lemma 5(3).** *The terminal basis  $b_1, \dots, b_n$  of SCFA satisfies with  $x' = x - \pi_n(x)$  the inequalities*

$$\|\widehat{b}_{i,x} - \widehat{b}_{i,x'}\| \leq \|\widehat{b}_{i,x}\| \quad \text{for } i = 1, \dots, n-1. \quad (3)$$

From  $\lambda(x') \geq 1/\max_{1 \leq i \leq n} \|\widehat{b}_{i,x'}\|$ ,  $\|a_n\| = \|\widehat{b}_{n,x'}\|^{-1}$  and  $\|\widehat{b}_{n-1,x'}\| = 0$  we see that

$$\frac{\|a_n\|}{\lambda(x')} \leq \max_{1 \leq i \leq n} \frac{\|\widehat{b}_{i,x'}\|}{\|\widehat{b}_{n,x'}\|} = \max_{1 \leq i \leq n-1} \{1, \|\widehat{b}_{i,x'}\| \|a_n\|\}.$$

From  $\|\widehat{b}_{i,x}\| \leq \epsilon$ , which in case  $x' \neq x$  holds upon termination of SCFA, and from the inequality  $\|\widehat{b}_{i,x'}\| \leq 2\|\widehat{b}_{i,x}\|$ , which follows from (3), we infer

$$\frac{\|a_n\|}{\lambda(x')} \leq \max\{1, 2\epsilon \|a_n\|\} \stackrel{(2)}{\leq} \max\{1, 2\epsilon 2^{n/2} \epsilon^{-1}\} = 2 \cdot 2^{n/2},$$

which finishes the proof.  $\square$

**Theorem 2.** *The dual bases  $b_1, \dots, b_n$  and  $a_1, \dots, a_n$ , occurring after the  $L^3$ -reduction step of SCFA, satisfy for  $i = 1, \dots, n-1$*

1.  $\|a_i\| \leq 1.5^{n-i} (\max_{i \leq j < n} \|\widehat{b}_{j,x}\|^{-1} + 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\})$ ,
2.  $\|b_i\| \leq 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\} \sum_{j=1}^i \prod_{\substack{k=1 \\ k \neq j}}^{n-1} \|\widehat{b}_{k,x}\|^{-1} + \sum_{j=1}^i \|\widehat{b}_{j,x}\|$ .

**Proof. 1.** Since SCFA did not terminate previously we know that  $\widehat{b}_{n,x} = 0$ , and thus  $\widehat{b}_{j,x} \neq 0$  holds for  $j = 1, \dots, n-1$ . Let  $\mu_{i,j}$  be the Gram-Schmidt coefficients of  $x, b_0, \dots, b_n$  and define the  $\nu_{i,j}$  by  $(\nu_{i,j})_{1 \leq i, j \leq n} := (\mu_{i,j})_{1 \leq i, j \leq n}^{-1}$ . We observe that

$$a_i = \sum_{j=i}^{n-1} \nu_{j,i} \frac{\widehat{b}_{j,x}}{\|\widehat{b}_{j,x}\|^2} + \nu_{n,i} a_n \quad (4)$$

holds for  $i = 1, \dots, n$ . In fact this formula implies

$$\begin{aligned}
\langle a_i, b_k \rangle &= \left\langle \sum_{j=i}^{n-1} \nu_{j,i} \frac{\widehat{b}_{j,x}}{\|\widehat{b}_{j,x}\|^2} + \nu_{n,i} a_n, \sum_{j=0}^k \mu_{k,j} \widehat{b}_{j,x} \right\rangle \\
&= \sum_{j=1}^{n-1} \nu_{j,i} \mu_{k,j} + \nu_{n,i} \langle a_n, b_k \rangle = \delta_{i,k} - \nu_{n,i} \mu_{k,n} - \nu_{n,i} \langle a_n, b_k \rangle = \delta_{i,k}.
\end{aligned}$$

The  $L^3$ -reduction step terminates with  $|\mu_{i,j}| \leq \frac{1}{2}$  for  $1 \leq j < i \leq n$ . Hence

$$|\nu_{i,j}| \leq 1.5^{i-j} \text{ for } 1 \leq j \leq i \leq n.$$

Now, equation (4) yields for  $i = 1, \dots, n$

$$\|a_i\|^2 \leq 1.5^{2(n-i)} \max_{i \leq j < n} \|\widehat{b}_{j,x}\|^{-2} + 1.5^{2(n-i)} \|a_n\|^2.$$

Using the inequality  $\|a_n\| \leq 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\}$  of Theorem 1 this proves the first claim:

$$\|a_i\| \leq 1.5^{n-i} \left( \max_{i \leq j < n} \|\widehat{b}_{j,x}\|^{-1} + 2^{n/2} \min\{\epsilon^{-1}, \lambda(x)\} \right).$$

**2.** We rewrite the equations (4) as

$$[a_1, \dots, a_n] = \left[ \frac{\widehat{b}_{1,x}}{\|\widehat{b}_{1,x}\|^2}, \dots, \frac{\widehat{b}_{n-1,x}}{\|\widehat{b}_{n-1,x}\|^2}, a_n \right] (\nu_{i,j})_{1 \leq i, j \leq n}.$$

Since the vectors  $\widehat{b}_{1,x}, \dots, \widehat{b}_{n-1,x}$  are pairwise orthogonal there is an orthogonal matrix  $U$ , i.e.  $U^{-1} = U^\top$ , such that

$$\begin{aligned}
&\left[ \frac{\widehat{b}_{1,x}}{\|\widehat{b}_{1,x}\|^2}, \dots, \frac{\widehat{b}_{n-1,x}}{\|\widehat{b}_{n-1,x}\|^2}, a_n \right] = \\
&U \begin{bmatrix} 1 & 0 & a'_{n,1} \\ & \ddots & \vdots \\ 0 & 1 & a'_{n,n-1} \\ 0 & \dots & 0 & a'_{n,n} \end{bmatrix} \begin{bmatrix} \|\widehat{b}_{1,x}\|^{-1} & & 0 & \vdots \\ & \ddots & & 0 \\ & & \|\widehat{b}_{n-1,x}\|^{-1} & \vdots \\ \dots & 0 & \dots & 1 \end{bmatrix},
\end{aligned}$$

with  $(a'_{n,1}, \dots, a'_{n,n})^\top := U a_n$  and  $a'_{n,n} = \|\widehat{b}_{1,x}\| \cdot \dots \cdot \|\widehat{b}_{n-1,x}\|$ . From the previous equations and  $[b_1, \dots, b_n]^\top = [a_1, \dots, a_n]^{-1}$  we see that

$$\begin{aligned}
&[b_1, \dots, b_n] = \\
&U \begin{bmatrix} 1 & & 0 & \vdots \\ & \ddots & & 0 \\ 0 & & 1 & \vdots \\ \bar{a}_{n,1} & \dots & \bar{a}_{n,n-1} & \bar{a}_{n,n} \end{bmatrix} \begin{bmatrix} \|\widehat{b}_{1,x}\| & & 0 & \vdots \\ & \ddots & & 0 \\ 0 & & \|\widehat{b}_{n-1,x}\| & \vdots \\ \dots & 0 & \dots & 1 \end{bmatrix} (\mu_{i,j})_{1 \leq i, j \leq n}^\top,
\end{aligned}$$

where  $\bar{a}_{n,n} := a'_{n,n}{}^{-1}$  and  $\bar{a}_{n,i} := -a'_{n,i}/a'_{n,n}$  for  $i < n$ . Since  $U$  is orthogonal  $\|b_i\|$  is the length of the  $i$ -th column vector of the cofactor of  $U$  in this matrix product. From  $\bar{a}_{n,n} = \|\widehat{b}_{1,x}\|^{-1} \cdot \dots \cdot \|\widehat{b}_{n-1,x}\|^{-1}$  and since the matrix  $(\mu_{i,j})_{1 \leq i,j \leq n}^\top$  is upper triangular with  $|\mu_{i,j}| \leq 1$  we have for  $i = 1, \dots, n-1$

$$\|b_i\|^2 \leq \|a_n\|^2 \sum_{j=1}^i \prod_{\substack{k=1 \\ k \neq j}}^{n-1} \|\widehat{b}_{k,x}\|^{-2} + \sum_{j=1}^i \|\widehat{b}_{j,x}\|^2 . \quad (5)$$

Now the claimed upper bound of  $\|b_i\|$  follows from the inequality  $\|a_n\| \leq 2^{n/2} \min\{\lambda(x), \epsilon^{-1}\}$  of Theorem 1.  $\square$

**Theorem 3.** For real input  $x = (x_1, \dots, x_{n-1}, 1)$  every vector  $(p_1, \dots, p_{n-1}, q) := b_1$  occuring after the  $L^3$ -reduction step of SCFA satisfies for  $i = 1, \dots, n-1$  :  $|x_i - p_i/q| \leq 2^{\frac{n+2}{4}} \sqrt{1+x_i^2} / |q|^{1+\frac{1}{n-1}}$ .

By a theorem of Dirichlet [Di1842] the upper bound  $\max_{1 \leq i \leq n} |x_i - p_i/q| \leq 1/|q|^{1+\frac{1}{n-1}}$  is best possible for diophantine approximations to  $x$  in that the term  $1/(n-1)$  can in general not be increased. SCFA loses at most the factor  $2^{(n+2)/4} \|x\|$  compared to this best general bound. This loss is due to the  $L^3$ -reduction. The factor  $2^{n/4}$  can be reduced to  $(1+\epsilon)^{n/4}$  with an arbitrarily small  $\epsilon > 0$  by using block reduced bases [Sc94].

**Proof.** For  $n = 2$  the sequence of rationals  $p_1/q$  occuring before a swap  $b_{n-1} \leftrightarrow b_n$  corresponds to the continued fraction expansion of  $x_1$ . Here we have the stronger inequality  $|x_1 - p_1/q| \leq 1/|q|^2$ . Now consider  $n \geq 3$ . From Theorem 2, the  $L^3$ -reducedness of  $\pi_{1,x}(b_1), \dots, \pi_{1,x}(b_{n-1})$  and  $\|\widehat{b}_{1,x}\| \leq 1$  we see that

$$\|b_1\| \leq 2^{n/2} \epsilon^{-1} \prod_{j=2}^{n-1} \|\widehat{b}_{j,x}\|^{-1} + \|\widehat{b}_{1,x}\| \leq 2^{n/2} \epsilon^{-1} \|\widehat{b}_{1,x}\|^{-n+2} 2^{(n-1)(n-2)/4} + 1 .$$

A look into the proofs of Theorem 1 and 2 shows that all inequalities, in particular inequality (2), hold with  $\epsilon^{-1}$  replaced by  $\|\widehat{b}_{1,x}\|^{-1}$ , and thus:

$$\|b_1\| \leq \|\widehat{b}_{1,x}\|^{-n+1} 2^{n/2} \cdot 2^{(n-1)(n-2)/4} + 1 \stackrel{n \geq 3}{\leq} \|\widehat{b}_{1,x}\|^{-n+1} 2^{(n-1)(n+2)/4} ,$$

$$\|\widehat{b}_{1,x}\| \leq \|b_1\|^{-\frac{1}{n-1}} 2^{(n+2)/4} \leq |q|^{-\frac{1}{n-1}} 2^{(n+2)/4} .$$

It can easily be seen that every vector  $b_1 = (p_1, \dots, p_{n-1}, q)$  satisfies for  $i = 1, \dots, n-1$

$$|x_i q - p_i| \leq \|\widehat{b}_{1,x}\| \sqrt{1+x_i^2} ,$$

see e.g. equations (18), (19) of [Ju92]. Now the claim follows from the previous upper bound on  $\|\widehat{b}_{1,x}\|$  and  $x_n = 1$ .  $\square$



**Theorem 4.** If SCFA outputs  $x' \neq x$  then  $a_n$  is, up to a factor  $2^{n/2+1}$ , a shortest almost relation for  $x$  in the following sense. If  $m \in \mathbb{Z}^n - 0$  satisfies  $|\langle x, m/\|m\| \rangle| < |\langle x, a_n/\|a_n\| \rangle|/2$  then  $\|m\| \geq \|a_n\| 2^{-n/2-1}$ .

**Proof.** Put  $\bar{x} := x - \langle x, m/\|m\| \rangle m/\|m\|$  and note that  $x' := x - \langle x, a_n/\|a_n\| \rangle a_n/\|a_n\|$ . Hence  $\|x - \bar{x}\| < \|x - x'\|/2$ . From [RS95], Theorem 6 we have  $\lambda(\bar{x}) \geq \epsilon^{-1}/2$  whereas  $\|a_n\| \leq 2^{n/2} \epsilon^{-1}$  holds by Theorem 1. Hence  $\|m\| \geq \lambda(\bar{x}) \geq \|a_n\| 2^{-n/2-1}$ .  $\square$

**Theorem 5.** For rational input  $x = (q_1, \dots, q_n)/q_n$  with integers  $q_1, \dots, q_n$ , SCFA performs at most  $\lceil \log_2 |q_n| \rceil$  swaps  $b_{n-1} \leftrightarrow b_n$ .

**Proof.** The vectors  $x, b_1, \dots, b_{n-1}$  before a swap  $b_{n-1} \leftrightarrow b_n$  are linearly independent, they generate a parallelepiped that has the non-zero volume:

$$\|x\| \prod_{j=1}^{n-1} \|\widehat{b}_{j,x}\| = \prod_{j=1}^{n-1} \|\widehat{b}_j\| \|\pi_n(x)\| = \|\widehat{b}_n\|^{-1} \|\pi_n(x)\| = \|a_n\| \|\pi_n(x)\|.$$

Hence  $\prod_{j=1}^{n-1} \|\widehat{b}_{j,x}\| = |\langle x, a_n \rangle| \|x\|^{-1}$ . This equation holds for the dual bases  $\bar{a}_1, \dots, \bar{a}_n, \bar{b}_1, \dots, \bar{b}_n$  before and the dual bases  $a_1, \dots, a_n, b_1, \dots, b_n$  after the swap. This yields

$$\frac{\|\widehat{b}_{n-1,x}\|}{\|\widehat{b}_{n-1,x}\|} = \frac{\prod_{j=1}^{n-1} \|\widehat{b}_{j,x}\|}{\prod_{j=1}^{n-1} \|\widehat{b}_{j,x}\|} = \frac{|\langle x, a_n \rangle|}{|\langle x, \bar{a}_n \rangle|}.$$

Using  $\|\widehat{b}_{n-1,x}\| = |\bar{\mu}_{n,n-1}| \|\widehat{b}_{n-1,x}\| \leq \frac{1}{2} \|\widehat{b}_{n-1,x}\|$  we see that

$$|\langle x, a_n \rangle| \leq \frac{1}{2} |\langle x, \bar{a}_n \rangle|.$$

Let  $t$  be the number of swaps  $b_{n-1} \leftrightarrow b_n$ . Since initially  $|\langle x, a_n \rangle| = 1$  we have upon termination of SCFA that  $q_n^{-1} \leq |\langle x, a_n \rangle| \leq 2^{-t}$ . This proves the desired bound  $t \leq \lceil \log_2 |q_n| \rceil$  on the number of swaps  $b_{n-1} \leftrightarrow b_n$ .  $\square$

**Running Time.** We refer to the models of computation introduced in section 2. Arithmetic operations are  $+$ ,  $-$ ,  $\cdot$ ,  $/$ ,  $\lceil \cdot \rceil$  (the nearest integer function) and  $<$  (comparison). In the floating point model we also use  $\sqrt{\cdot}$  (square root).

*Exact Real Arithmetic.* For real input  $x \in \mathbb{R}^n$  SCFA performs  $O(n^4 (n + |\log \epsilon|))$  arithmetic operations on real numbers and  $O(n^2 (n + |\log \epsilon|))$  many swaps  $b_{k-1} \leftrightarrow b_k$  with  $2 \leq k \leq n$ . This follows from the analysis of the HJLS-algorithm [HJLS89]. The algorithm either uses Gram-Schmidt orthogonalization via the  $\mu_{i,j}$  and  $\|\widehat{b}_{j,x}\|^2$  or Givens Rotation with square roots via the  $\tau_{i,j}$ .

*Exact Integer Arithmetic.* For rational  $x = (q_1, \dots, q_n)/q_n \in \mathbb{Q}^n$  SCFA performs at most  $O(n^4 (n + |\log \epsilon|))$  arithmetic operations on integers of bit length  $O(n + \max_{1 \leq i \leq n} |\log q_i| + |\log \epsilon|)$ . Arithmetic steps use the coordinates of the

vectors  $b_i$ ,  $a_i$  and the numerators and denominators of the rational numbers  $\mu_{i,j}$ ,  $\|\widehat{b}_{j,x}\|^2$ . The algorithm uses Gram–Schmidt orthogonalization. The claimed upper bound on the bit length of all integers follows by adjusting the analysis of the  $L^3$ –algorithm in [LLL82] to our algorithm.

## 5 Numerical Stability of the Floating Point Algorithm

The algorithm is given a rational input  $x = (q_1, \dots, q_n)/q_n \in \mathbb{Q}^n$ . Orthonormalization of  $x, b_1, \dots, b_n$  is done by Givens Rotation via the floating point numbers  $\tau_{i,j} := \mu_{i,j} \|\widehat{b}_{j,x}\|$ . We study floating point errors of the  $\tau_{i,j}$  and the correctness of swaps  $b_{k-1} \leftrightarrow b_k$ .

Every arithmetic operation  $+$ ,  $-$ ,  $\cdot$ ,  $/$ ,  $\sqrt{\phantom{x}}$ ,  $\lceil \phantom{x} \rceil$  generates a floating point error when its result is rounded to the nearest floating point number. Let  $t'$  denote the floating point value of a real number  $t$  with (floating point) *error*  $t - t'$  and *relative error*  $(1 - t'/t)$ . Let  $r$  denote the number of precision bits of the floating point arithmetic and  $2^{-r}$  the maximal relative error. We use IEEE 754 double precision format with  $r = 53$ .

**Givens Rotation.** The  $n \times (n+1)$ –matrix  $B := (b_{i,j}) = [x, b_1, \dots, b_n]$  has a unique decomposition  $B = U \cdot L^\top$  where  $L$  is a lower triangular matrix and  $U$  is an orthogonal  $n \times n$ –matrix. Hence  $L = (\tau_{i,j})_{\substack{0 \leq i \leq n \\ 1 \leq j \leq n}}$  and

$U = \left[ \begin{array}{c} x \\ \frac{\widehat{b}_{1,x}}{\|\widehat{b}_{1,x}\|}, \dots, \frac{\widehat{b}_{n-1,x}}{\|\widehat{b}_{n-1,x}\|} \end{array} \right]$ . Since  $U$  is orthogonal we have  $L^\top = U^\top B$ .  $U^\top = U^{-1}$  is product of *elementary rotations* (ER)  $G_{i,j}$  with  $1 \leq j < i \leq n$ . If  $\overline{B} = (\overline{b}_{i,j}) = [x, \overline{b}_1, \dots, \overline{b}_n]$  denotes the product of  $B$  with all previous ER then  $\overline{B} \mapsto G_{i,j} \overline{B}$  puts  $\overline{b}_{i,j}$  to zero by transforming the column vectors  $\overline{b}_i, \overline{b}_j$ .

**Floating Point Errors.** Let  $|B| := \max_{0 \leq i \leq n} \|b_i\|$ . Note that  $|B| = |\overline{B}|$  holds since the  $G_{i,j}$  are orthogonal. Multiplying  $B$  with  $G_{i,j}$  yields an error  $\|\overline{b}_j - \overline{b}_j'\|$ ,  $\|\overline{b}_i - \overline{b}_i'\| \leq 7n2^{-r}|B|$ , see pp. 131–139 in [Wi65]. This also holds for the multiplication by several  $G_{i,j}$  with pairwise disjoint sets  $\{i, j\}$ . Following Gentleman [Ge75] we can distribute the  $(n-2)(n-1)/2$  many ER  $G_{i,j}$  for  $U$  into  $2n-3$  stages, each containing pairwise disjoint  $G_{i,j}$ . This way the error of the  $\tau_{i,j}$  produced by Givens Rotation is at most  $7(2n-3)2^{-r}|B|$ , see [Ge75], [H95].

We call a swap  $b_{k-1} \leftrightarrow b_k$  *good* if it decreases  $\tau_{k-1,k-1}$ , i.e. if the swap condition  $\tau_{k-1,k-1} > (\tau_{k,k}^2 + \tau_{k,k-1}^2)^{1/2}$  holds before a swap.

**Theorem 6.** *If  $\frac{3}{4}\tau_{k-1,k-1}^2 > \tau_{k,k}^2 + \tau_{k,k-1}^2$  holds for the rounded  $\tau$ –values of a Givens Rotation for the actual basis  $B$  and  $195n2^{-r}|B| \leq \tau_{k-1,k-1}$  the swap  $b_{k-1} \leftrightarrow b_k$  is good.*

**Proof.** Above, we have shown that the error of the matrix  $[\bar{b}_1, \dots, \bar{b}_n] := (\tau_{i,j})^\top$  satisfies  $\|\bar{b}_{k-1} - \bar{b}'_{k-1}\|, \|\bar{b}_k - \bar{b}'_k\| \leq 7(2n-3)2^{-r}|B|$  and thus

$$\tau_{k-1,k-1} - (\tau_{k,k}^2 + \tau_{k,k-1}^2)^{1/2} \geq (1 - \sqrt{\frac{3}{4}})\tau_{k-1,k-1} + \sqrt{\frac{3}{4}}\tau'_{k-1,k-1} - (\tau_{k,k-1}^{\prime 2} + \tau_{k,k}^{\prime 2})^{1/2} - (\sqrt{\frac{3}{4}} + 1)7(2n-3)2^{-r}|B|.$$

By assumption  $\sqrt{\frac{3}{4}}\tau'_{k-1,k-1} - \sqrt{\tau_{k,k-1}^{\prime 2} + \tau_{k,k}^{\prime 2}} + 2^{-r+1} > 0$  holds where  $2^{-r+1}$  bounds the errors in evaluating the swap condition. These inequalities imply  $\tau_{k-1,k-1} > (\tau_{k,k}^2 + \tau_{k,k-1}^2)^{1/2}$  since  $14(\sqrt{\frac{3}{4}} + 1)/(1 - \sqrt{\frac{3}{4}}) < 195$ .  $\square$

By the argument of Theorem 6 the inequalities  $\frac{1}{2}\tau_{k-1,k-1}^2 > \tau_{k,k}^2 + \tau_{k,k-1}^2$ ,  $165n2^{-r}|B| \leq \tau_{k-1,k-1}$  imply that the swap condition  $\frac{3}{4}\tau_{k-1,k-1}^{\prime 2} \geq \tau_{k,k-1}^{\prime 2} + \tau_{k,k}^{\prime 2}$  holds for the rounded  $\tau$ -values. (Here we use that  $14(\sqrt{\frac{3}{4}} + 1)/(\sqrt{\frac{3}{4}} - \sqrt{\frac{1}{2}}) < 165$ ). Therefore the  $L^3$ -reduction step generates a reasonably good approximation of an  $L^3$ -reduced basis.

**Good Swaps for SCFA.** From inequality (5) and since the basis  $B$  before a swap is reduced in size we have  $|B| \leq \sqrt{n}2^{n/2}\epsilon^{-1}(\sum_{i=1}^n q_i^2)^{1/2}$ . Hence by Theorem 6 a swap  $b_{k-1} \leftrightarrow b_k$  of SCFA is good if  $\tau_{k-1,k-1} > \epsilon$  and

$$195n^22^{n/2}\epsilon^{-2}2^{-r}\max_i |q_i| \leq 1.$$

E.g. for  $r = 53$ ,  $\epsilon^{-1} = 16$ ,  $|q_i| \leq 2^{20}$  this inequality holds up to  $n = 18$ . Since  $\tau_{k-1,k-1} > \epsilon$  holds for all but a few exceptional swaps we see that SCFA is stable for  $\epsilon^{-1} = 16$ ,  $\max_i |q_i| \leq 2^{20}$  up to dimension 18. In fact our experiments show that the stability of SCFA is even much better.

**Experimental Results.** We distinguish the two cases of termination:

- an exchange step  $b_{n-1} \leftrightarrow b_n$  results in  $\tau_{n,n} \neq 0$  and in an integer relation  $a_n$  for  $x$
- $\max_{i=1,\dots,n} \|\widehat{b}_{i,x}\| < \epsilon$  yields a nearby point  $x' \neq x$ .

The last column of our table reports the number of occurrences of the second case. In the first case we check that the vector  $a_n$  satisfies  $\langle a_n, x \rangle = 0$ .

Each line of the table shows the results for 10 input vectors  $x := (q_1, \dots, q_n)/q_n \in \mathbb{Q}^n$  where the  $q_i \in_R [0, 2^Q - 1]$  are random integers. With  $Q = 45$  we almost exhaust the 53 precision bits of double precision floating point. We see that the average length of the output vector  $a_n$  increases at most linearly with  $\epsilon^{-1}$ . For  $n = 10$  we have  $16 < \lambda(x) \leq 58.45$  for all 10 inputs.

$n$	$\epsilon^{-1}$	$Q$	$av. len$	$av. dist$	$av. time$	$\# x'$
10	4	45	7.53	1.987e+7	0.18	10
10	8	45	8.71	20.76	0.28	10
10	16	45	21.46	4.7	0.47	10
10	24	45	20.40	0.54	0.37	9
10	32	45	30.23	0.0	0.37	3
10	64	45	29.95	0.0	0.47	0
20	2	45	4.17	2.256e+4	3.07	10
20	3	45	6.27	1.124e+3	4.26	10
20	4	45	8.68	0.04	5.41	4
20	6	45	7.92	0.0	5.34	0
modified SCFA						
20	16	45	7.92	0.0	5.41	0
40	16	45	7.87	0.0	75.01	0
60	16	45	10.83	0.0	328.14	0
80	16	45	7.45	0.0	958.77	0
100	16	45	6.32	0.0	2268.43	0

$n, \epsilon$  : parameters of SCFA  
 $Q$  : the above parameter  
 $av. len$  : average length of the output vector  $a_n$  per 10 inputs  
 $av. dist$  : average distance from  $x$  to  $x'$  per 10 inputs  
 $av. time$  : average time in seconds per 10 inputs on a HP 715/50 with 62 MIPS  
 $\# x'$  : number of nearby points  $x' \neq x$  for 10 inputs

In dimensions  $n \geq 20$  and  $Q = 45$  SCFA runs out of numerical precision. To improve the stability we perform before each size-reduction step and before each swap  $b_{k-1} \leftrightarrow b_k$  a fresh Givens Rotation for the actual vectors  $x, b_1, \dots, b_n$ , so that Theorem 6 applies. With this modification, SCFA becomes numerically stable up to dimension 100 and  $Q = 45$ .

Reasonable running times demonstrate not only the efficiency but also the stability of SCFA, they show that good swaps  $b_{k-1} \leftrightarrow b_k$  prevail. Then the output  $a_n$  is most likely correct since a faulty output  $a_n$  means a bad swap  $b_{n-1} \leftrightarrow b_n$ . The fact that SCFA always finds integer relations  $a_n$  as expected shows its stability.

## References

- [Berg80] G. BERGMAN: Notes on Ferguson and Forcade's Generalized Euclidean Algorithm. TR, Department of Mathematics, University of California, Berkeley, CA, 1980.
- [Bern71] L. BERNSTEIN: The Jacobi-Perron Algorithm, Lecture Notes in Mathematics 207, Berlin-Heidelberg-New York (1971), pp. 1-161.
- [BBP96] D. BAILEY, P. BORWEIN, S. PLOUFFE: On the Rapid Computation of Various Polylogarithmic Constants, Technical Report, Simon Frazer University, Burnaby, B. C., Canada (1996).

- [Di1842] G.L. DIRICHLET: Verallgemeinerung eines Satzes aus der Lehre von den Kettenbrüchen nebst einigen Anwendungen auf die Theorie der Zahlen, Bericht über die zur Bekanntmachung geeigneten Verhandlungen der Königlich Preussischen Akademie der Wissenschaften zu Berlin (1842), pp. 93–95.
- [FB92] H.R.P. FERGUSON and D.H. BAILEY: A Polynomial Time, Numerically Stable Integer Relation Algorithm. RNR Technical Report RNR-91-032, NASA Ames Research Center, Moffett Field, CA (1992).
- [FF79] H. FERGUSON and R. FORCADE: Generalization of the Euclidean Algorithm for Real Numbers to all Dimensions Higher than Two, Bull. Amer. Math. Soc., (New Series) 1 (1979), pp. 912–914.
- [Ge75] W.M. GENTLEMAN: Error Analysis of QR Decomposition by Givens Transformations. Linear Algebra and its Applications, Vol. 10, pp. 189–197, 1975.
- [GL89] G.H. GOLUB and C.F. VAN LOAN: Matrix Computations. The Johns Hopkins University Press, London (1989).
- [H95] C. HECKLER: Automatische Parallelisierung und parallele Gitterbasenreduktion. Ph.D. Thesis, University of Saarbrücken, 1995.
- [HT93] C. HECKLER and L. THIELE: A Parallel Lattice Basis Reduction for Mesh-connected Processor Arrays and Parallel Complexity. Proceedings of the 5th Symposium on Parallel and Distributed Processing, Dallas (1993).
- [HJLS89] J. HASTAD, B. JUST, J.C. LAGARIAS and C.P. SCHNORR: Polynomial Time Algorithms for Finding Integer Relations among Real Numbers. SIAM J. Comput., Vol. 18, No. 5 (1989), pp. 859–881.
- [Ja1868] C.G.J. JACOBI: Allgemeine Theorie der Kettenbruchähnlichen Algorithmen, J. Reine Angew. Math. 69 (1868), pp. 29–64.
- [Jo93] A. JOUX: A Fast Parallel Lattice Basis Reduction Algorithm. Proceedings of the 2nd Gauss Symposium, Munich (1993).
- [Ju92] B. JUST: Generalizing the Continued Fraction Algorithm to Arbitrary Dimensions. SIAM J. Comput., Vol. 21, No. 5 (1992), pp. 909–926.
- [LLL82] A.K. LENSTRA, H.W. LENSTRA, JR. and L. LOVÁSZ: Factoring Polynomials with Rational Coefficients. Math. Ann. 21 (1982), pp. 515–534.
- [Pe1907] O. PERRON: Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus. Math. Ann. 64 (1907), pp. 1–76.
- [RS95] C. RÖSSNER and C.P. SCHNORR: Computation of Highly Regular Nearby Points. Proceedings of the 3rd Israel Symposium on Theory of Computing and Systems, Tel Aviv (1995).
- [Sc94] C.P. SCHNORR: Block Reduced Lattice and Successive Minima. Combinatorics, Probability and Computing 3 (1994), pp. 507–522.
- [Sz70] G. SZEKERES: Multidimensional Continued Fractions. Ann. Univ. Sci. Budapest, Eötvös Sect. Math. 13 (1970), pp. 113–140.
- [Wi65] J.H. WILKINSON: The Algebraic Eigenvalue Problem. Oxford University Press (1965).