to appear in *Theoretical Computer Science* **172** (February 1997)

Finding a shortest vector in a two-dimensional lattice modulo m

Günter Rote*

September 16, 1996

Abstract

We find the shortest non-zero vector in the lattice of all integer multiples of the vector (a, b) modulo m, for given integers $0 < a, b < m$. We reduce the problem to the computation of a Minkowski-reduced basis for a planar lattice and thereby show that the problem can be solved in $O(\log m(\log\log m)^2)$ bit operations.

1 Introduction

Let $0 < a, b < m$ be integers. We define the *lattice* $L_m((a, b))$ *generated by the vector* (a, b) *modulo* m as the set of vectors $\{\, (ta \bmod m, tb \bmod m) \mid 0 \le t < m \,\}$, which forms a group under componentwise addition modulo m. For clarification, we emphasize that $x \bmod d$ denotes the unique integer in the interval $[0, m-1]$ which is congruent to x. Thus we regard $L_m((a, b))$ as a subset of $[0, m-1]^2$. By the length $\|(a, b)\|$ of a vector (a, b) we will always mean its Euclidean length $\sqrt{a^2 + b^2}$, no matter whether (a, b) is considered as a vector modulo m or as a usual vector in the plane.

Lempel and Paz [1994] considered the computation of the shortest non-zero vector in $L_m((a, b))$. They pointed out that this problem is different from the computation of the shortest non-zero vector in a planar lattice in the usual sense, for which there is a classical algorithm of Gauß, which takes $O(\log m)$ arithmetic steps. Hence they developed a completely new algorithm taking also $O(\log m)$ arithmetic steps.

The purpose of this paper to show that this problem can be reduced to the classical case very easily, and therefore it is possible to apply results from the literature, in particular the lattice reduction algorithm of Yap [1992], which, in terms of bit complexity, takes only $O(\log\log m)$ times as many steps as one integer multiplication.

We will briefly review the basics of reduced lattice bases to make the paper self-contained. The two crucial geometric observations, which are stated in lemmas 2 and 3, are completely elementary, and the algorithm directly follows from them.

2 Preliminaries

A two-dimensional *lattice* is the set of all integer linear combinations of two linearly independent vectors x and y. These two vectors are said to *generate* $L(x, y)$ and they are called a *basis* of the lattice.

*Institut für Mathematik, Technische Universität Graz, Steyrergasse 30, A-8010 Graz, Austria; electronic mail: rote@opt.math.tu-graz.ac.at

It is obvious that the set $L_m((a,b))$ is essentially the same as the lattice $L := L((a,b), (m,0),(0,m))$ generated by the vectors (a,b), $(m,0)$ and $(0,m)$: If we restrict L to $[0,m-1]^2$, we get $L_m((a,b))$, and by adding to $L_m((a,b))$ all integer multiples of the vectors $(m,0)$ and $(0,m)$ we obtain L.

Lempel and Paz [1994] showed by an example that the shortest non-zero vector in $L_m((a,b))$ may differ from the shortest non-zero vector in L. However, we have the following relation.

Lemma 1 *The shortest non-zero vector in* $L_m((a,b))$ *is the shortest non-zero vector in* L *inside the first quadrant* $Q = \{\,(x,y) \mid x,y \geq 0\,\}$.

Proof: If $a+b > m$, then the vector $(a',b') = (m,0)+(0,m)-(a,b) \in L$ fulfills $a'+b' < m$. Thus either the point (a,b) or the point (a',b') lies in the first quadrant Q and below or on the line $x+y=m$. If follows that the shortest non-zero vector in $L \cap Q$ has length less than m, and therefore it must belong to $L \cap [0,m-1]^2 = L_m((a,b))$. ■

3 Finding a basis

Before we look for shortest vectors in the lattice $L = L((a,b),(m,0),(0,m))$ we have to find a basis, i. e., a set of only two vectors f and g which generate L. It is well-known how this can be done in general: Let u, v, and w be three vectors which generate some two-dimensional lattice. Since they are linearly dependent and rational, we can write

$$t_1 u + t_2 v + t_3 w = 0,$$

for some integers t_1, t_2, t_3 with $\gcd(t_1,t_2,t_3) = 1$. (In our case, we can directly set $t_1 = m$, $t_2 = -a$, $t_3 = -b$ and divide these numbers by their greatest common divisor.) We extend the vector (t_1,t_2,t_3) to an integer matrix

$$A = \begin{pmatrix} t_1 & \alpha_1 & \beta_1 \\ t_2 & \alpha_2 & \beta_2 \\ t_3 & \alpha_3 & \beta_3 \end{pmatrix}$$

with $\det A = 1$. We show below how to find such a matrix. Now a basis consists of the vectors $f = \alpha_1 u + \alpha_2 v + \alpha_3 w$ and $g = \beta_1 u + \beta_2 v + \beta_3 w$, or in matrix notation,

$$(u,v,w)A = (0,f,g),$$

where the vectors u, v, w, f, and g are written as column vectors. Since f and g are given by integer linear combinations of u, v, and w, it is clear that $L(f,g) \subseteq L(u,v,w)$. On the other hand, we have

$$(u,v,w) = (0,f,g)A^{-1}$$

with an integer matrix A^{-1}. Thus u, v, and w can also be expressed as integer linear combinations of f and g, and $L(u,v,w) \subseteq L(f,g)$. Therefore f,g is really a basis of $L(u,v,w)$.

It remains to select numbers α_i and β_i so that the matrix A has determinant 1, see for example Theorem 14.2.3 of Hua [1982, p. 376]. The extended Euclidean algorithm yields integers α_1, α_2, β_3, and γ with

$$\begin{aligned} d := \gcd(t_1,t_2) &= \alpha_2 t_1 - \alpha_1 t_2, && (1) \\ 1 = \gcd(d,t_3) &= \beta_3 d - \gamma t_3. && (2) \end{aligned}$$

We set $\alpha_3 = 0$, $\beta_1 = t_1/d \cdot \gamma$, and $\beta_2 = t_2/d \cdot \gamma$. The equation $\det A = 1$ can be checked easily using (1) and (2), for example by expanding the third row of A.

4 Basis reduction

A *reduced basis* for a lattice (in the sense of Minkowski) consists of two vectors x_1, x_2 with the following properties.

1. x_1 is a shortest non-zero lattice vector.

2. x_2 is a shortest vector among the lattice vectors which are not parallel to x_1.

It follows that x_1 and x_2 form a basis of the lattice.

For the purpose of the proofs of the following two lemmas we may assume without loss of generality that $\|x_2\| \geq \|x_1\| = 1$. For the inner product we have then the bound $-1/2 \leq \langle x_1, x_2 \rangle \leq 1/2$, because otherwise one of the vectors $x_2 \pm x_1$ would be shorter than x_2. By possibly switching the sign of x_2 we can additionally achieve $\langle x_1, x_2 \rangle \geq 0$.

Lemma 2 *Let x_1 and x_2 be a reduced basis, and assume without loss of generality that $\langle x_1, x_2 \rangle \geq 0$. Then the following holds.*

1. *x_1 and $-x_1$ are two shortest non-zero vectors.*

2. *x_2 and $-x_2$ are two shortest vectors among the those vectors which are not parallel to x_1.*

3. *$x_3 := x_2 - x_1$ and $-x_3$ together with x_2 and $-x_2$ are four shortest vectors among the those vectors which are not parallel to x_1.*

Proof: Only the third statement does not follow immediately from the definition. Any lattice point x which is not parallel to x_1 can be uniquely represented as $x = \alpha x_1 + \beta x_2$ with two integers α and $\beta \neq 0$. Let $w = x_2 - \langle x_1, x_2 \rangle \cdot x_1$ be the component of x_2 which is orthogonal to x_1. Then x can be written as a sum of two orthogonal components:

$$x = \beta w + \langle x, x_1 \rangle x_1$$

For a fixed β, the length of x is therefore minimized when $|\langle x, x_1 \rangle|$ is minimized. Let us first consider the points x with $\beta = 1$. The set of all values $\langle x, x_1 \rangle = \langle x_2 + \alpha x_1, x_1 \rangle$ for $\alpha \in \mathbb{Z}$ forms an infinite arithmetic progression with increment 1. The two values which are smallest in absolute value are clearly $\langle x_2, x_1 \rangle \in [0, 1/2]$ and $\langle x_3, x_1 \rangle = \langle x_2, x_1 \rangle - 1 \in [-1, -1/2]$. Thus the vectors x_2 and x_3, together with their inverses $-x_2$ and $-x_3$, form a set of four candidates for the four shortest vectors among the lattice vectors not parallel to x_1. We have finished the proof if we can show that all vectors x with $|\beta| \geq 2$ are longer than x_3. The distance of the line $\{ \lambda x_1 + \beta x_2 \mid \lambda \in \mathbb{R} \}$ from the origin is $|\beta| \cdot \|w\| \geq 2\|w\|$. This is a lower bound for the length of any vector x on this line. By the Pythagorean theorem we have

$$\|w\|^2 = \|x_2\|^2 - \langle x_2, x_1 \rangle^2 = \|x_2\|^2 - \left(\frac{\langle x_2, x_1 \rangle}{\|x_2\|} \right)^2 \|x_2\|^2 \geq \|x_2\|^2 (1 - 1/4),$$

and thus $2\,\|w\| \geq \sqrt{3}\,\|x_2\|$ is a lower bound for $\|x\|$. On the other hand, $\|x_3\|^2 = \|w\|^2 + \langle x_3, x_1 \rangle^2 \leq \|x_2\|^2 + 1 \leq 2\,\|x_2\|^2$, and therefore, with $\|x_3\| \leq \sqrt{2}\,\|x_2\|$, x_3 is shorter than x. ■

Lemma 3 *Under the assumptions of the previous lemma, the vectors* $x_1, x_2, x_3, -x_1, -x_2, -x_3$ *form a hexagon surrounding the origin, and the angle between any two successive vectors is at most* $90°$.

Proof: This is immediate for the angles between x_1 and x_2 and between x_3 and $-x_1$ (and their inverse vectors). Only for x_2 and x_3 we need a little computation.

$$\|x_2\| \cdot \|x_3\| \cdot \cos(\angle x_2 x_3) = \langle x_3, x_2 \rangle = \langle x_2 - x_1, x_2 \rangle = \langle x_2, x_2 \rangle - \langle x_1, x_2 \rangle \geq 1 - 1/2 > 0.$$

(In fact, it can be shown that this angle is at most $60°$.) ■

5 The Algorithm

Now it is clear how to proceed. By the previous lemma, one of the six vectors is contained in the first quadrant Q. If neither x_1 nor $-x_1$ is in Q, no vector parallel to x_1 is in Q and it suffices to check the remaining vectors. The four shortest ones among them are x_2, $-x_2$, x_3, and $-x_3$, and by just checking them in this order we find the shortest vector in Q. Here is a precise summary of the algorithm.

1. Find a basis for the lattice L generated by the vectors (a, b), $(m, 0)$, and $(0, m)$.

2. Find a reduced basis x_1, x_2 for this lattice L. Ensure that $\langle x_1, x_2 \rangle \geq 0$ by changing the sign of x_2 if necessary.

3. (a) If x_1 or $-x_1$ lies in the first quadrant Q, then this is the shortest vector in $L_m((a, b))$.
   (b) Otherwise, if x_2 or $-x_2$ belongs to Q, then this is the shortest vector in $L_m((a, b))$.
   (c) Otherwise, either $x_2 - x_1$ or $-(x_2 - x_1)$ belongs to Q, and this is the shortest vector in $L_m((a, b))$.

Step 1 involves four computations of a greatest common divisor of two integers which are at most m, for which there is an algorithm taking $O(\log\log m \cdot M(\log m))$ bit operations, where $M(n)$ denotes the bit complexity of multiplying two n-bit integers [Schönhage 1971]. Step 2 is is a generalization of the greatest common divisor computation. Extending the technique of Schönhage, Yap [1992] showed that a reduced basis can be computed in the same time bound. Therefore, Step 2 can also be carried out in $O(\log\log m \cdot M(\log m))$ time. The rest is just a constant number of arithmetic operations on integers of absolute value at most m. On a random access machine (RAM) model with logarithmic cost (see Aho, Hopcroft, and Ullman [1974, p. 12] for a definition of this model), we have $M(n) = O(n \log n)$ [Schönhage 1980, Theorem 7.1]. Therefore, the overall bit complexity of our procedure is $O(\log m (\log\log m)^2)$.

6 Generalizations. Higher dimensions

We will briefly discuss the extension of our approach to the higher-dimensional version of the problem, as well as other possible generalizations.

Lemmas 2 and 3 provide an explicit list of linear combinations of the basis vectors of a reduced basis, which is guaranteed to contain the shortest lattice vector in a given quadrant. One may ask whether such a list exists in higher dimensions, or even for other cones than an orthant.

If this is true in higher dimensions, then the problem of the shortest vector in a d-dimensional modular lattice can be reduced to finding a reduced lattice basis in a d-dimensional lattice, a well-studied problem which can be solved satisfactorily in theory (see [Kannan 1987] and [Helfrich 1985]) as well as in practice [Finke and Pohst 1985]. In higher dimensions, a *reduced* lattice basis can be taken in any of several different meanings, see for example [Vallée 1989] for an overview.

It seems likely that analogs of Lemmas 2 and 3 hold in any dimension, although the required list of linear combinations will probably grow at least exponentially as the dimension increases.

7 References

A. V. Aho, J. Hopcroft, and J. D. Ullman [1974]
The Design and Analysis of Computer Algorithms, Addison-Wesley, 1974.

Ulrich Fincke and M. Pohst [1985]
Improved methods for calculating vectors of short length in a lattice, including a complexity analysis, *Math. Comp.* **44**, 463–471.

Bettina Helfrich [1985]
Algorithms to construct Minkowski and Hermite reduced bases, *Theoret. Comput. Sci.* **41**, 125–139.

Hua L.-K. [1982]
Introduction to Number Theory, Springer-Verlag, 1982.

R. Kannan [1987]
Minkowski's convex body theorem and integer programming, *Math. Oper. Res.* **12**, 415–440.

M. Lempel and A. Paz [1994]
An algorithm for finding a shortest vector in a two-dimensional modular lattice, *Theoret. Comput. Sci.* **125**, 229–241.

Arnold Schönhage [1971]
Schnelle Berechnung von Kettenbruchentwicklungen, *Acta Informatica* **1**, 139–144.

Arnold Schönhage [1980]
Storage modification machines, *SIAM J. Comput.* **9**, 490–508.

Brigitte Vallée [1989]
La réduction des réseaux. Autour de l'algorithme de Lenstra, Lenstra, Lovász, *RAIRO Inform. Théor. Appl.* **23**, 345–376. English translation: A central problem in the algorithmic geometry of numbers: Lattice reduction. Around the algorithm of Lenstra, Lovász, Lovász, *CWI Quarterly* **3** (no. 2), 95–120 (1990).

Chee K. Yap [1992]
Fast unimodular reduction: Planar integer lattices, In: *Proc. 33rd Ann. IEEE Sympos. Found. Comput. Sci. (FOCS)*, 1992, pp. 437–446.