Note on shortest and nearest lattice vectors

Martin Henk

Fachbereich Mathematik, Sekr. 6-1 Technische Universität Berlin Straße des 17. Juni 136 D-10623 Berlin Germany henk@math.tu-berlin.de

We show that with respect to a certain class of norms the so called shortest lattice vector problem is polynomial-time Turing (Cook) reducible to the nearest lattice vector problem. This gives a little more insight in the relationship of these two fundamental problems in the computational geometry of numbers.

Key words: Computational Geometry, shortest lattice vector, nearest lattice vector, polar lattice.

1 Introduction

Throughout this paper let \mathbb{R}^n be the real *n*-dimensional vector space equipped with a norm $f_K(\cdot)$, where K is the gauge-body of the norm, i.e.

$$K = \{ x \in \mathbb{R}^n : f_K(x) \le 1 \}.$$

K is a centrally symmetric convex body with nonempty interior and $f_K(\cdot)$ is also called the distance function of K because $f_K(x) = \min\{\rho \in \mathbb{R}^{\geq 0} : x \in \rho K\}$. The Euclidean norm is denoted by $f_B(\cdot)$, where B is the n-dimensional unit ball, and the associated inner product is denoted by $\langle \cdot, \cdot \rangle$. Finally, we denote by C the cube with edge length 2 and center 0, and thus $f_C(\cdot)$ denotes the maximum norm. As usual we denote by $\lceil x \rceil$ the smallest integer not less than $x \in \mathbb{R}$.

Let $b^1, \ldots, b^n \in \mathbb{Q}^n$ be *n* linearly independent vectors. The set

$$\Lambda = \left\{ x \in \mathbb{Q}^n : x = \sum_{i=1}^n z_i b^i, \ z_i \in \mathbb{Z}, \ 1 \le i \le n \right\}$$

Preprint submitted to Elsevier Preprint

30 September 2007

is called the lattice generated by the basis b^1, \ldots, b^n .

Now, the shortest lattice vector problem with respect to a norm $f_K(\cdot)$ – SVP_K — is the following task (cf. [5]):

SVP_K: Let $\Lambda \subset \mathbb{Q}^n$ be a lattice given by a basis b^1, \ldots, b^n . Find a lattice vector $b \in \Lambda \setminus \{0\}$ with minimal distance to 0, i.e. $f_K(b) = \min\{f_K(w) : w \in \Lambda \setminus \{0\}\}.$

The length of a shortest nonzero vector of a lattice Λ with respect to a norm $f_K(\cdot)$ is denoted by $\lambda_K(\Lambda)$.

The nearest vector problem with respect to the norm $f_K(\cdot) - \text{NVP}_K$ — is in a certain sense the inhomogeneous counterpart to SVP_K :

NVP_K: Let $\Lambda \subset \mathbb{Q}^n$ be a lattice given by a basis b^1, \ldots, b^n and let $v \in \mathbb{Q}^n$. Find a lattice vector $c \in \Lambda$ with minimal distance to v, i.e. $f_K(c-v) = \min\{f_K(w-v) : w \in \Lambda\}.$

Observe that in the SVP_K we are looking for a nonzero lattice vector, whereas in the NVP_K the zero vector is a solution for all sufficiently small vectors v. Geometrically speaking the NVP_K is the task to find a lattice point c such that the given point v is contained in the honeycomb $c + H_K(\Lambda)$, where $H_K(\Lambda)$ is given by:

$$H_K(\Lambda) = \{ x \in \mathbb{R}^n : f_K(x) \le f_K(x-w), \quad \forall w \in \Lambda \}.$$

 $H_K(\Lambda)$ is a centrally symmetric ray set, i.e. if $x \in H_K(\Lambda)$, then $\rho x \in H_K(\Lambda)$ for all ρ with $-1 \leq \rho \leq 1$. In the Euclidean case, but not in general, $H_K(\Lambda)$ is a convex set. Moreover, it is easy to see that the inradius of $H_K(\Lambda)$ with respect to $f_K(\cdot)$ is one half of the length of a shortest lattice vector. Analogously, the circumradius $\mu_K(\Lambda)$ of a honeycomb is equal to the so called inhomogeneous minimum of Λ and $f_K(\cdot)$ (cf. [6]), which is defined as the maximal distance of a point to the lattice, i.e. $\mu_K(\Lambda) = \max_{y \in \mathbb{R}^n} \min_{w \in \Lambda} f_K(w - y) = \max\{f_K(x) : x \in H_K(\Lambda)\}$. Thus

$$\frac{\lambda_K(\Lambda)}{2}K \subseteq H_K(\Lambda) \subseteq \mu_K(\Lambda)K.$$
(1)

It is known that the nearest vector problem is \mathcal{NP} -hard for the norms $f_B(\cdot)$ and $f_C(\cdot)$, see VAN EMDE BOAS [3] and KANNAN [7]. Probably the shortest vector problem is also \mathcal{NP} -hard, but up to now this has only been established with respect to the maximum norm (cf. [3]). The purpose of this note is to prove that for a certain class of norms SVP_K is not harder than NVP_K : **Theorem 1.1** Let $\Lambda \subset \mathbb{Q}^n$ be a lattice given by a basis b^1, \ldots, b^n and let $f_K(\cdot)$ be a norm with the property that

$$f_C(x) \le f_K(x) \le f_B(x), \quad x \in \mathbb{R}^n.$$
(2)

With polynomially many calls to a subroutine solving NVP_K and polynomial additional time we can solve SVP_K .

At a first sight this class of norms appears to very special, but for example all *p*-norms $|x|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ are part of this class for $p \ge 2$.

We assume that the function $f_K(\cdot)$ is computable in polynomial time. The input size of our problem is given by the input size of the vectors b^1, \ldots, b^n . For detailed information about complexity and numerical computation we refer to the book [5] and for notation concerning lattices we refer to [6].

Finally, we note that (2) is equivalent to $B \subseteq K \subseteq C$, and thus implies that

$$\frac{1}{\sqrt{n}}f_B(x) \le f_K(x) \le f_B(x),\tag{3}$$

$$f_K(e^i) = 1, \quad 1 \le i \le n, \tag{4}$$

where e^i denotes the *i*-th unit vector.

2 Proof of Theorem 1.1

Now we state an algorithm which reduces the problem SVP_K to NVP_K . Since the algorithm works inductively the input is given by an *m*-dimensional lattice Λ embedded in \mathbb{R}^n , where we assume m > 1. Otherwise it is trivial to compute a shortest nonzero lattice vector. The linear hull of Λ is denoted by $lin(\Lambda)$.

Procedure SVP_{K} -by_NVP_K:

Input: An *m*-dimensional lattice Λ generated by the basis $b^1 \ldots, b^m \in \mathbb{Q}^n$ and a norm $f_K(\cdot)$ on \mathbb{R}^n such that (2) is satisfied.

Output: A shortest nonzero vector b of the lattice Λ with respect to $f_K(\cdot)$.

(i) With respect to the Euclidean norm find an "almost" shortest nonzero lattice vector b^* in the polar lattice

$$\Lambda^* = \{ z \in \ln(\Lambda) : \langle z, w \rangle \in \mathbb{Z}, \forall w \in \Lambda \}$$

of Λ by calls of the subroutine $\mathrm{NVP}_K,$ i.e. find a primitive vector $b^* \in \Lambda^*$ such that

$$\lambda_B(\Lambda^*) \ge \frac{f_B(b^*)}{2n}.$$
(5)

(ii) Find a basis $\bar{b}^1, \ldots, \bar{b}^m$ of Λ such that

$$\langle \bar{b}^i, b^* \rangle = 0, \ 1 \le i \le m-1, \quad \text{and} \quad \langle \bar{b}^m, b^* \rangle = 1.$$
 (6)

- (iii) Let $H_i = \{x \in \text{lin}(\Lambda) : \langle b^*, x \rangle = i\}, i \in \mathbb{Z}$. For $1 \leq i \leq \lfloor 2n^{3/2} \cdot m \rfloor$ find a shortest lattice vector $u^{m,i}$ in the affine hyperplane H_i using the subroutine NVP_K.
- (iv) Let u^m be a lattice vector of minimal length among the vectors $u^{m,i}$.
- (v) Find a shortest lattice vector u^{m-1} in the plane H_0 by applying the procedure SVP_K-by_NVP_K to the (m-1)-dimensional lattice Λ^{m-1} generated by the vectors $\bar{b}^1, \ldots, \bar{b}^{m-1}$ and the norm $f_K(\cdot)$.
- (vi) Let b be the shorter one of u^m and u^{m-1} . Then $f_K(b) = \lambda_K(\Lambda)$.

Proof of Theorem 1.1 Without loss of generality we may assume $\Lambda \subseteq \mathbb{Z}^n$. First we prove the correctness of the algorithm. Obviously, a shortest lattice vector of Λ is contained in $\bigcup_{i=0}^{\infty} H_i$. It remains to show that it suffices to consider the planes H_i , $0 \leq i \leq \lfloor 2n^{3/2}m \rfloor$. By a result of BANASZCZYK [1] (see also BOURGAIN & MILMAN [2]) we have $\lambda_B(\Lambda) \cdot \lambda_B(\Lambda^*) \leq m$. Since $B \subseteq K$ by (2) we have $\lambda_B(\Lambda) \geq \lambda_K(\Lambda)$ and thus

$$\lambda_K(\Lambda) \le \frac{m}{\lambda_B(\Lambda^*)}.$$

On account of (5) we obtain

$$\lambda_K(\Lambda) \le \frac{2n \cdot m}{f_B(b^*)}.\tag{7}$$

On the other hand we have for each vector $y \in H_i$ that $f_B(y) \ge i/f_B(b^*)$ and so (cf. (3))

$$f_K(y) \ge \frac{i}{\sqrt{n}f_B(b^*)}.$$

Hence for $i > \lfloor 2n^{3/2} \cdot m \rfloor$ the length of a shortest lattice vector in a plane H_i is greater than $\lambda_K(\Lambda)$.

In the sequel we show how the single steps of the algorithm can be done.

Step (i):

Let b^{m+1}, \ldots, b^n be an orthogonal basis of the orthogonal complement of $lin(\Lambda)$. Such a basis can be found in polynomial time via the GRAM-SCHMIDT orthogonalization (cf. [5]). Let B be the matrix with columns b^1, \ldots, b^n . Then the first m columns of the matrix $(B^T)^{-1}$ form a basis of Λ^* . Let the columns of $(B^T)^{-1}$ be denoted by b^{1*}, \ldots, b^{n*} and let $\tilde{\Lambda}^*$ be the n-dimensional lattice generated by $b^{1*}, \ldots, b^{m*}, \sigma b^{m+1*}, \ldots, \sigma b^{n*}$ with

$$\sigma = \left[\frac{2n}{\min\{f_B(b^{j_*}) : m+1 \le j \le n\}} f_B(b^{1_*})\right] + 1$$
(8)

In what follows we construct a vector $b^* \in \tilde{\Lambda}^* \setminus \{0\}$ such that (5) holds with $\tilde{\Lambda}^*$ instead of Λ^* , i.e.

$$\lambda_B(\tilde{\Lambda}^*) \ge \frac{f_B(b^*)}{2n}.$$
(9)

Then by the choice of σ we will see that b^* belongs to Λ^* and thus the vector satisfies (5).

Since the parallelepiped P^* spanned by $b^{1*}, \ldots, b^{m*}, \sigma b^{m+1*}, \ldots, \sigma b^{n*}$ generates a lattice tiling with respect to $\tilde{\Lambda}^*$ the width $\omega(P^*)$ of P^* is a lower bound for $\lambda_B(\tilde{\Lambda}^*)$. Now

$$\omega(P^*) = \min\left\{1/f_B(b^1), \dots, 1/f_B(b^m), \sigma/f_B(b^{m+1}), \dots, \sigma/f_B(b^n)\right\}$$

and so

$$\lambda_B(\tilde{\Lambda}^*) \ge \omega(P) \ge \gamma := \min\left\{\frac{1}{\lceil f_B(b^i) \rceil}, \quad 1 \le i \le n\right\}.$$

With $\nu_0 = \gamma / \lceil 2\sqrt{n} \rceil$ we obtain for $1 \le i \le n$ (cf. (4))

$$f_K(\nu_0 e^i) = \nu_0 \le \frac{\lambda_B(\Lambda^*)}{2\sqrt{n}} \le \frac{\lambda_K(\Lambda^*)}{2}.$$
(10)

Thus the vectors $\nu_0 e^i$ belong to the honeycomb $H_K(\tilde{\Lambda}^*)$ (cf. (1)) and the origin is the unique nearest lattice vector to $\nu_0 e^i$. Moreover for $\nu_1 = \sum_{i=1}^m [f_B(b^{i_*})] + \sum_{i=m+1}^n [f_B(\sigma b^{i_*})]$ we obtain (cf. [6])

$$f_K(\nu_1 e_i) = \nu_1 > \frac{1}{2} \left(\sum_{i=1}^m f_K(b^{i_*}) + \sum_{i=m+1}^n f_K(\sigma b^{i_*}) \right) \ge \mu_K(\tilde{\Lambda}^*).$$
(11)

Hence the vectors $\nu_1 e^i$ are not contained in $H_K(\tilde{\Lambda}^*)$. On account of (1) and since $H_K(\tilde{\Lambda}^*)$ is a ray set the output of the subroutine NVP_K with input Λ and νe^i is a nonzero lattice vector for $\nu \geq \nu_1$. So by applying the subroutine NVP_K to the sequence of points $2^k \nu_0 e^i$, $k = 0, \ldots$, we can find after at most $\lceil \log_2(\nu_1) - \log_2(\nu_0) \rceil$ calls of NVP_K a positive scalar ϵ_i with $\nu_0 \leq \epsilon_i \leq \nu_1$ and a nonzero lattice vector v^{i_*} such that

$$\epsilon_i e^i \in H_K(\tilde{\Lambda}^*)$$
 and $2\epsilon_i e^i \in v^{i_*} + H_K(\tilde{\Lambda}^*)$.

The last relation implies $f_K(2\epsilon_i e^i - v^{i_*}) \leq f_K(2\epsilon_i e^i)$ and thus $f_K(v^{i_*}) \leq 4\epsilon_i$. Now let $\epsilon_k = \min\{\epsilon_i : 1 \leq i \leq n\}$ and let $b^* = v^{k_*}$. In the sequel we show that b^* satisfies (9). For abbreviation we write ϵ instead of ϵ_k . On account of (3) we get

$$f_B(b^*) \le 4\epsilon \sqrt{n}.\tag{12}$$

 $H_K(\tilde{\Lambda}^*)$ is a centrally symmetric ray set and thus $\pm \epsilon \cdot e^i \in H_K(\tilde{\Lambda}^*)$. Hence (cf. (4))

$$f_B(\epsilon e^i) = \epsilon = f_K(\epsilon e^i) \le f_K(\epsilon e^i - u^*) \le f_B(\epsilon e^i - u^*)$$
(13)

for all lattice vectors $u^* \in \tilde{\Lambda}^*$. That means that the vectors $\pm \epsilon e^i$, $1 \leq i \leq n$, are contained in the honeycomb $H_B(\tilde{\Lambda}^*)$ which is a convex set. So the width of the cross polytope with vertices $\pm \epsilon \cdot e^i$ is a lower bound for $\lambda_B(\tilde{\Lambda}^*)$:

$$\lambda_B(\tilde{\Lambda}^*) \ge \frac{2\epsilon}{\sqrt{n}}.$$

Together with (12) we obtain

$$\lambda_B(\tilde{\Lambda}^*) \geq \frac{f_B(b^*)}{2n}$$

Now suppose $b^* \notin \Lambda^*$. Then by the definition of $\tilde{\Lambda}^*$ and σ we have

$$f_B(b^*) \ge \sigma \min\{f_B(b^{j_*}) : l+1 \le j \le n\} > 2nf_B(b^{l_*}) \ge 2n \cdot \lambda_B(\Lambda^*)$$

which contradicts the choice of b^* . Obviously, we may assume that b^* is primitive.

Step (ii):

In order to find a basis of Λ such that (6) is satisfied we first construct a basis $\bar{b}^{1*}, \ldots, \bar{b}^{m*}$ of Λ^* containing b^* . Furthermore, without loss of generality we may assume that the matrix M with columns $b^*, b^{2*}, \ldots, b^{m*}$ has rank m. Let A be the integer $(m \times m)$ -matrix such that $M = B^* \cdot A$, where B^* is the matrix with columns b^{1*}, \ldots, b^{m*} . It is well known that the HERMITE normal form of a A can be computed in polynomial time (cf. [4], [8] or [10]) and thus we can find an unimodular matrix U and an integer upper triangle matrix T with $A = U \cdot T$. Hence $M = (B^* \cdot U) \cdot T$ and the columns $\bar{b}^{1*}, \ldots, \bar{b}^{m*}$ of $B^* \cdot U$ form a basis of Λ^* . Since b^* is primitive b^* is the first column vector of $B^* \cdot U$. Now, let b^{m+1*}, \ldots, b^{n*} be an orthogonal basis of the orthogonal complement of $\ln(\Lambda)$ and \bar{B}^* be the matrix with columns $\bar{b}^{1*}, \ldots, \bar{b}^{m*}, b^{m+1*}, \ldots, b^{n*}$. Then the first m columns of $(\bar{B}^T)^{-1}$ form a basis of Λ such that (6) is satisfied.

Step (iii):

Obviously, $H_i \cap \Lambda = \{x \in \Lambda : x = \sum_{j=1}^{m-1} z_j \bar{b}^j + i \bar{b}^m\}$. Hence, for $i \ge 1$ the shortest vector problem in the plane H_i is the task to find a lattice vector of Λ^{m-1} which is nearest to $-i \bar{b}^m$. That can be done by applying the procedure NVP_K to the input vector $-i b^m$ and a suitable *n*-dimensional lattice $\hat{\Lambda}$. For example, let g^m, \ldots, g^n be an orthogonal basis of the orthogonal complement of $\ln(\Lambda^{m-1})$ and let

$$\chi = \lceil 2\sqrt{n} \rceil \lceil 2n^{3/2} \cdot m \rceil \left\lceil \frac{f_B(b^m)}{\min\{f_B(g^j) : m \le j \le n\}} \right\rceil + 1$$

We claim that a nearest lattice vector of the lattice $\hat{\Lambda}$ generated by $\bar{b}^1, \ldots, \bar{b}^{m-1}, \chi g^m, \ldots, \chi g^n$ to $-ib^m$ belongs to Λ^{m-1} . Suppose the opposite and let $z_1b^1 + \cdots + z_{m-1}b^{m-1} + z_m\chi g^m + \cdots + z_n\chi g^n$ be a nearest lattice vector to $-ib^n$ with some $z_j \neq 0$, say z_n , for $m \leq j \leq n$. Then $f_K(z_1b^1 + \cdots + z_{m-1}b^{m-1} + z_m\chi g^m + \cdots + z_n\chi g^n - ib^l) \leq if_K(b^m)$ and on account of (3) we get

$$\frac{|z_n|\chi f_B(g^n)}{\sqrt{n}} - if_K(b^m)$$

$$\leq \frac{f_B(z_1b^1 + \dots + z_{m-1}b^{m-1} + z_m\chi g^m + \dots + z_n\chi g^n)}{\sqrt{n}} - if_K(b^m)$$

$$\leq f_K(z_1b^1 + \dots + z_{m-1}b^{m-1} + z_m\chi g^m + \dots + z_n\chi g^n - ib^m) \leq if_K(b^m).$$

It follows $\chi \leq 2\sqrt{n}if_B(b^m)/f_B(g^n) \leq \lceil 2\sqrt{n} \rceil \lceil 2n^{3/2} \cdot m \rceil \cdot f_B(b^m)/f_B(g^n)$ which contradicts the choice of χ .

To analyze the encoding length of the numbers arising in the procedure let $\langle A \rangle$ be the input size of the algorithm and let Λ^i be the *i*-dimensional lattice constructed in the (m-i)-th call of the procedure SVP_{K} -by_NVP_K. Furthermore, let b^{*_i} be the "almost" shortest lattice vector of the appropriate dual

lattices Λ^{*_i} constructed in step (i). Then $\det(\Lambda^{j-1}) = \det(\Lambda^j) f_B(b^{*_j})$ and by (5) and MINKOWSKI's convex body theorem (cf. [5])

$$\det(\Lambda^{j-1}) = \det(\Lambda^j) f_B(b^{*_j}) \le \det(\Lambda^j) 2n\lambda_B(\Lambda^{*_j})$$
$$\le \det(\Lambda^j) 2n\sqrt{j} \det(\Lambda^{*_j})^{1/j} \le \det(\Lambda^j) 2n \cdot \sqrt{j}$$

So det (Λ^{j-1}) < det $(\Lambda^m)(2n)^{2m}$. Before we call the procedure SVP_K-by_NVP_K we can make an LLL-reduction and obtain a basis $c^1, \ldots, c^j \in \mathbb{Z}^n$ of Λ^j with (cf. [5])

$$f_B(c^1)\cdots f_B(c^j) \leq 2^{j^2} \det(\Lambda^j).$$

Hence, for each vector c^{j} we have the bound

$$f_B(c^j) \le 2^{m^2} (2n)^{2m} \det(\Lambda^m).$$

This implies that we can always find a basis of the lattice Λ^j whose input size is bounded by $O(\langle A \rangle^3)$. Finally, it is easy to check that the numbers arising in the steps (i)–(iv) of the procedure SVP_{K} -by_NVP_K are bounded by a polynomial in the input size of the given basis at the beginning of the execution of the steps (i)–(iv) (cf. [5]).

3 Remarks

The crucial point for the special choice of the norms given by (2) is relation (13). For example, if $f_K(\cdot)$ is the 1-norm, i.e., K is the cross polytope with edge length $\sqrt{2}$ then, in general, it is not true that $f_K(\epsilon e^i - u^*) \leq f_B(\epsilon e^i - u^*)$. Hence we do not know whether $\pm \epsilon e^i$, $1 \leq i \leq n$, are contained in $H_B(\tilde{\Lambda}^*)$ and we can not show that b^* is an "almost" shortest lattice vector as required in step (i) (cf. proof of Theorem 1.1). On the other hand if K is an arbitrary centrally symmetric convex body and if we have oracles solving NVP_K and NVP_B, then SVP_K can be solved in oracle polynomial time, because by the subroutine solving NVP_B we can easily find such an "almost" shortest lattice vector. However, we believe that SVP_K is polynomial reducible to NVP_K at least for any *p*-norm.

The above algorithm works also for any norm $f_{\bar{K}}(\cdot)$ for which an affine transformation L exists such that $B \subset L\bar{K} \subset C$ because $f_{L\bar{K}}(x) = f_{\bar{K}}(L^{-1}x)$.

Finally, we remark that in the Euclidean case the problem to find a KORKIN-ZOLOTAREV reduced basis of a lattice is polynomial reducible to the shortest lattice vector problem SVP_B (cf. [9]). By Theorem 1.1 we obtain that this problem can also be reduced to NVP_B in polynomial time.

References

- W. Banaszczyk, New bounds in some transference theorems in the geometry of numbers, Math. Ann. 296 (1993), no. 4, 625–635.
- [2] J. Bourgain and V.D. Milman, Sections euclidiennes et volume des corps symétriques convexes dans ℝⁿ, C.R. Acad. Sci. Paris Sér. I Math. **300** (1985), 435–438.
- [3] P. van Emde Boas, Another NP-complete partition problem and the complexity of computing short vectors in a lattice, Report 81-04, Mathematical Institute, University of Amsterdam (1981).
- [4] M.A. Frumkin, An algorithm for the reduction of a matrix of integers to triangular form with power complexity of the computations, Ekonomika i Matematicheskie Metody 12 (1976), 173–178 (Russian).
- [5] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer, Berlin, 1988.
- [6] P.M. Gruber and C.G. Lekkerkerker, *Geometry of Numbers*, 2.nd ed., North-Holland, Amsterdam, 1987.
- [7] R. Kannan, Minkowski's convex body theorem ad integer programming, Math. Operations Research 12 (1987), no. 3, 415–440.
- [8] R. Kannan and A. Bachem, Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix, SIAM Journal on Computing 8 (1979), 499–507.
- [9] L. Lovász, An Algorithmic Theory of Numbers, Graphs and Convexity, (CBMS-NSF Regional Conference Series in Applied Mathematics 50), SIAM, Philadelphia, Pennsylvania, 1986.
- [10] M. Pohst and H. Zassenhaus, Algorithmic Algebraic Number Theory, Cambridge Univ. Press, Cambridge, 1989.