A PROCEDURE FOR DETERMINING ALGEBRAIC INTEGERS OF GIVEN NORM

U. Fincke and M. Pohst

Mathematisches Institut
Universität Düsseldorf
Universitätsstr. 1
4 Düsseldorf, West Germany

1. Introduction:

Let F be an algebraic number field of degree n. Let s denote the number of real, $2t$ the number of complex conjugates of F ordered in the usual way such that

$$F^{(1)},\ldots,F^{(s)} \subseteq \mathbb{R}\,,\quad F^{(s+t+j)} = \overline{F^{(s+j)}} \subseteq \mathbb{C}\quad (1 \leq j \leq t).$$

Let R be a subring of the ring of algebraic integers in F which contains $\mathbb{Z}$ and is a free $\mathbb{Z}$-module of rank n:

$$R = \mathbb{Z}\,w_1 + \ldots + \mathbb{Z}\,w_n\,.$$

One of the key problems in algebraic number theory is the solution of norm equations: For given $K \in \mathbb{Z}^{\geq 2}$ we want to determine $\beta \in R$ with

$$(1)\quad |N(\beta)| = \Big|\prod_{i=1}^{n} \beta^{(i)}\Big| = K\,,$$

where $\beta^{(1)},\ldots,\beta^{(n)}$ denote the conjugates of $\beta = \beta^{(1)}$.

This problem for example arises in all class number and class group computations.

Our method of solving (1) bases on K. Mahler's "ceilings" (introduced in [5]) which allow to transform the product in (1) into a positive definite quadratic form depending on n real parameters $\underline{\lambda}=(\lambda_1,\ldots,\lambda_n)$. A solution of (1) uniquely determines $\underline{\lambda}$. By appropriate methods of mathematical optimization we obtain only a small number of vectors $\underline{\lambda}$ which possibly yield a solution of (1). For each such $\underline{\lambda}$ it is easily tested, whether it corresponds to a solution of (1).

2. Relaxation and separation of the problem.

Each $\beta \in R$ has a presentation

(2) $\beta = x_1 w_1 + \ldots + x_n w_n \ (x_j \in \mathbb{Z},\ 1 \leq j \leq n)$.

Writing $< \underline{w}^{(i)}, \underline{x} > := \sum_{j=1}^{n} x_j w_j^{(i)} = \beta^{(i)} \ (1 \leq i \leq n)$ for abbreviation,

(1) is equivalent to

(3) $|N(\beta)| = |\prod_{i=1}^{s} < \underline{w}^{(i)}, \underline{x} > | \prod_{i=s+1}^{s+t} | < \underline{w}^{(i)}, \underline{x} > |^2 = K$

$(\underline{x} \in \mathbb{Z}^n, \ \underline{x} \neq \underline{0})$.

If (3) has a solution, then - as a consequence of the following lemma - there is also a solution $\underline{x}$ of (3) for which the coordinates x_i of $\underline{x}$ are bounded. Because of $x_i \in \mathbb{Z}$ $(1 \leq i \leq n)$ there are only finitely many possibilities for $\underline{x}$.

Lemma 1. Let $\varepsilon_1, \ldots, \varepsilon_{s+t-1}$ be a system of independent units of R and (3) be solvable. Then there exists

$\beta = x_1 w_1 + \ldots + x_n w_n \in R$ and $L_i, U_i \in \mathbb{R}^{>0} (1 \leq i \leq n)$ satisfying (1) and

(4) $L_i \leq |\beta^{(i)}| \leq U_i$

$$\text{for } L_i := \exp\left(\frac{1}{n} \log K - \frac{1}{2} \sum_{j=1}^{s+t-1} |\log |\varepsilon_j^{(i)}||\right),$$

$$U_i := \exp\left(\frac{1}{n} \log K + \frac{1}{2} \sum_{j=1}^{s+t-1} |\log |\varepsilon_j^{(i)}||\right)$$

$(1 \leq i \leq n)$.

A proof of Lemma 1 is given in [1]. We easily derive bounds for the coefficients x_j $(1 \leq j \leq n)$ of β, for example, via a dual basis of R.

Since the amount of computations necessary to solve (3) strongly depends on the size of L_i, U_i $(1 \leq i \leq n)$, $\varepsilon_1, \ldots, \varepsilon_{s+t-1}$ are to be chosen such that

$$\sum_{j=1}^{s+t-1} |\log |\varepsilon_j^{(i)}|| \quad (1 \leq i \leq n)$$

respectively their product, become small. Thus it would be of advantage, if $\varepsilon_1, \ldots, \varepsilon_{s+t-1}$ were fundamental units of R. However, it is very time consuming to show that $s + t - 1$ independent units are already

fundamental units. Hence, in general we will be satisfied with s+t-1 independent units which are fundamental with high probability. They can be computed by the methods of [6],[7]. In the sequel we always assume that $\varepsilon_1,\ldots,\varepsilon_{s+t-1}$ are independent units of R and L_i,U_i ($1 \le i \le n$) the constants defined in Lemma 1.
The next lemma transforms the problem of solving (3) into one which turns out to be easier to handle.

Lemma 2. Let (3) be solvable. Then there exists $\beta = x_1w_1+\ldots+x_nw_n = \langle \underline{w},\underline{x}\rangle \in R$ satisfying (1) and (4), and $\underline{\lambda} \in (\mathbb{R}^{>0})^{s+t}$ subject to

$$(5)\quad \prod_{i=1}^{s} \lambda_i \prod_{i=s+1}^{s+t} \lambda_i^2 = 1,$$

$$(6)\quad \sum_{i=1}^{s} \lambda_i |\langle \underline{w}^{(i)},\underline{x}\rangle|^2 + 2 \sum_{i=s+1}^{s+t} \lambda_i |\langle \underline{w}^{(i)},\underline{x}\rangle|^2 = nK^{2/n},$$

$$(7)\quad U_i^{-2}K^{2/n} \le \lambda_i \le L_i^{-2} K^{2/n} \quad (1 \le i \le s+t).$$

Proof. According to Lemma 1 there exists $\beta = \langle \underline{w},\underline{x}\rangle \in R$ satisfying $|N(\beta)| = K$ and (4). For this algebraic integer β we define:

$$\lambda_i := |\langle \underline{w}^{(i)},\underline{x}\rangle|^{-2} K^{2/n} \in \mathbb{R}^{>0} \quad (1 \le i \le n).$$

Then (5), (6), (7) are an immediate consequence of (3), (4). □

Remark. The transformation of (3) into (5), (6) follows an idea of K. Mahler [5].

Instead of solving (3) directly we try to solve (6) subject to (4), (5), (7) instead. This turns out to be easier, since we can prove that it suffices to consider the quadratic form in (6) for very few vectors $\underline{\lambda} = (\lambda_1,\ldots,\lambda_{s+t})$.

Theorem 1. Let (3) be solvable and λ, R_i, S_i ($1 \le i \le n$) defined by:

(8) $\lambda \in \mathbb{R}^{>1}$ is a (unique) zero of

$$(1-h(\lambda))\,\lambda^{h(\lambda)} + h(\lambda)\lambda^{h(\lambda)-1} - \left(1+\frac{1}{K}\right)^{2/n} = 0$$

$$\text{for } h(\lambda) := \frac{\lambda}{\lambda-1} - \frac{1}{\log \lambda},$$

(9) $R_i := \lfloor \frac{2}{\log \lambda} (\frac{1}{n} \log K - \log U_i) \rfloor$,

$S_i := \lceil \frac{2}{\log \lambda} (\frac{1}{n} \log K - \log L_i) \rceil = -R_i (1 \le i \le n)$.

Then there exists $\beta = \langle \underline{w}, \underline{x} \rangle \in R$ with $|N(\beta)| = K$ satisfying

(10) $\sum_{i=1}^{n} \lambda^{r_i} |\langle \underline{w}^{(i)}, \underline{x} \rangle|^2 \le n(K+1)^{2/n}$,

where $r_1, \ldots, r_n$ are rational integers subject to

(11a) $\sum_{i=1}^{n} r_i = 0$,

(11b) $R_i \le r_i \le S_i \quad (1 \le i \le n)$,

(11c) $r_{s+t+i} - r_{s+i} \in \{0,1\} \ (1 \le i \le t)$ and

$\#\{i | r_{s+t+i} - r_{s+i} = 1, \ 1 \le i \le t\} \in \{0,1\}$.

Proof. It is easily seen that $h(\lambda) := \frac{\lambda}{\lambda-1} - \frac{1}{\log \lambda}$ is in $(0,1)$ for $\lambda > 1$. Then for $x \in (0,1)$, $\lambda > 1$ the function $f(x,\lambda) := (1-x)\lambda^x + x\lambda^{x-1}$ is strictly increasing in λ for fixed x and concave in x for fixed λ with a maximum for $x_\lambda := \frac{\lambda}{\lambda-1} - \frac{1}{\log \lambda}$. Therefore $f(x_\lambda, \lambda)$ is strictly increasing in λ. Because of $f(x_{1+}, 1+) = 1$ we obtain a unique zero λ of $f(x_\lambda, \lambda) - (1 + \frac{1}{K})^{2/n} = 0$, hence (8).

We define $y_i := K^{-2/n} |\beta^{(i)}|^2$ $(1 \le i \le n)$ for $\beta \in R$ satisfying the conditions of Lemma 2. We note that $\prod_{i=1}^{n} y_i = 1$ because of $|N(\beta)| = K$. We can represent y_i by λ in the form

$$y_i = \lambda^{-\tilde{r}_i + \tilde{\varepsilon}_i} \ (0 \le \tilde{\varepsilon}_j < 1, \tilde{r}_i \in \mathbb{Z}, \ \sum_{i=1}^{n} \tilde{r}_i = \sum_{i=1}^{n} \tilde{\varepsilon}_i = 0,$$

$$1 \le j \le n-1, \ 1 \le i \le n).$$

By subtracting 1 from $\tilde{\varepsilon}_i$ for suitable indices $i \in \{1, \ldots, n-1\}$ and changing $\tilde{r}_i, \tilde{\varepsilon}_n, \tilde{r}_n$ correspondingly, we obtain $r_1, \ldots, r_n \in \mathbb{Z}$, $a \in (0,1)$ and $\varepsilon_1, \ldots, \varepsilon_n \in [a-1, a]$ such that $y_i = \lambda^{-r_i + \varepsilon_i}$ $(1 \le i \le n)$ and (11a), (11c) are sasitsfied. Then (11b), (9) are an easy consequence of (7).

Finally we prove (10). Namely,

$$\sum_{i=1}^{n} \lambda^{r_i} y_i = \sum_{i=1}^{n} \lambda^{\varepsilon_i} \leq \sum_{i=1}^{n} ((1-(a-\varepsilon_i))\lambda^a + (a-\varepsilon_i)\lambda^{a-1})$$

$$\leq n(1-h(\lambda))\lambda^{h(\lambda)} + n\, h(\lambda)\lambda^{h(\lambda)-1}$$

$$= n(1 + \frac{1}{K})^{2/n} .$$

The first inequality follows from the convexity of the exponential function, the second by $f(x,\lambda) \leq f(x_\lambda,\lambda)$, and the last equation by (8).

□

For the computations it is of advantage to split the matrix of coefficients of the quadratic form in (10):

Lemma 3. The quadratic form in (10) satisfies

(12) $\sum_{i=1}^{n} \lambda^{r_i} |\langle \underline{w}^{(i)}, \underline{x} \rangle|^2 = \underline{x} U D_{\underline{\lambda}}^2 U^{tr} \underline{x}^{tr}$ for

$$U^{tr} := \begin{pmatrix} \underline{w}^{(1)} \\ \vdots \\ \underline{w}^{(s)} \\ \operatorname{Re} \underline{w}^{(s+1)} \\ \operatorname{Im} \underline{w}^{(s+1)} \\ \vdots \\ \operatorname{Re} \underline{w}^{(s+t)} \\ \operatorname{Im} \underline{w}^{(s+t)} \end{pmatrix} \quad \text{and}$$

$\underline{\lambda} := (v_1,\ldots,v_s,z_1,\ldots,z_{2t})$, $D_{\underline{\lambda}} = \operatorname{diag}(v_1,\ldots,v_s,z_1,\ldots,z_{2t})$,

where

$v_i := \lambda^{r_i/2} (1 \leq i \leq s)$, $z_{2i-1} = z_{2i} := (\lambda^{r_{s+i}} + \lambda^{r_{s+t+i}})^{1/2}$

$(1 \leq i \leq t)$

and

$\operatorname{Re} \underline{w}^{(s+i)} := (\operatorname{Re} w_1^{(s+i)},\ldots,\operatorname{Re} w_n^{(s+i)})$,

$\operatorname{Im} \underline{w}^{(s+i)} := (\operatorname{Im} w_1^{(s+i)},\ldots,\operatorname{Im} w_n^{(s+i)})$ $(1 \leq i \leq t)$.

Proof. By straightforward computation.

3. Worst case analysis.

The number of quadratic forms (12) which must be considered is roughly bounded by

(13) $(t+1)\,(3-\frac{4\log K}{n\log\lambda}+\frac{4\log B}{\log\lambda})^{s+t-1}$

in case $U_i \le B$ $(1 \le i \le n)$. This follows from the conditions for $r_i (1 \le i \le n)$ of Theorem 1. We note that (13) yields a reasonable bound only in case λ is not close to 1. This imposes conditions for the size of n and K.

For each quadratic form $Q_{\underline{\lambda}}$ obtained we need to determine the set

$$S_{\underline{\lambda}} := \{\underline{x} \in \mathbb{Z}^n \setminus \{\underline{0}\} \mid nK^{2/n} \le \underline{x}UD_{\underline{\lambda}}^2 U^{tr}\underline{x}^{tr} \le n(K+1)^{2/n}\}$$

according (12) and (10) Also by (10) $S_{\underline{\lambda}}$ can contain only vectors $\underline{x}$ with $|N(<\underline{w},\underline{x}>)| \le K+1$ and the probability for $|N(<\underline{w},\underline{x}>)| = K+1$ $(\underline{x} \in S_{\underline{\lambda}})$ is practically 0. Hence, $S_{\underline{\lambda}}$ contains only very few points which are then good candidates for solving (3).

Theorem 2. Denote $D_o := \mathrm{diag}(1,..,1,\sqrt{2},...,\sqrt{2}) = D_{\underline{\lambda}_o}$ and $V^{(o)} = (v_{ij}^{(o)})_{1\le i,j\le n} := (D_o U^{tr})^{-1}$. Let $v, B \in \mathbb{R}^{>o}$ such that $|v_{ij}^{(o)}| \le v$ $(1 \le i,j \le n)$ and $U_i \le B$ $(1 \le i \le n)$. Then for sufficiently large B the number of arithmetic operations of our procedure is roughly bounded by

(14) $(t+1)\,(3-\frac{4\log K}{n\log\lambda}+\frac{4\log B}{\log\lambda})^{s+t-1}\ \frac{1}{2}(2^{\frac{n+1}{2}}\,B\,nv\,\lambda^{\frac{1}{2}}(1+\frac{1}{K})^{\frac{1}{n}}+1)^{\frac{n}{2}}\,2(n+1)n.$

Proof. The first two factors come from the number (13) of quadratic forms to be considered. The last two factors contain the number of necessary operations for the transition from one quadratic form to the next, for the reduction algorithm applied to each quadratic form and, finally for the enumeration procedure.

Let $V^{(\underline{\lambda})} := (D_{\underline{\lambda}} U^{tr})^{-1}$ with $D_{\underline{\lambda}}$ defined as in Lemma 3. We denote the rows of $V^{(\underline{\lambda})}$ by $\underline{v}_1^{(\underline{\lambda})},...,\underline{v}_n^{(\underline{\lambda})}$ and set $T := \max\{S_i \mid 1 \le i \le n\}$ for abbreviation. Applying the reduction algorithm of [4] to the rows of an arbitrary regular matrix $V \in \mathbb{R}^{n,n}$ we get its row-reduced version $\tilde{V}$. Because of $|v_{ij}^{(o)}| \le v, |v_{ij}^{(\underline{\lambda})}| \le v\lambda^{T/2}$ $(1 \le i,j \le n)$ and (1.12) of [4] we obtain

(15) $\| \tilde{\underline{v}}_i^{(o)} \|^2 \leq 2^{n-1}\, n v^2 \quad (1 \leq i \leq n)$

(16) $\| \tilde{\underline{v}}_i^{(\underline{\lambda})} \|^2 \leq 2^{n-1}\, n \lambda^T v^2 \quad (1 \leq i \leq n)$,

where $\| . \|$ denotes the Euclidean norm in $\mathbb{R}^n$.
Furthermore, Proposition 1.6 from [4] yields

(17) $|\det V^{(o)}|^2 \leq \prod_{i=1}^{n} \| \tilde{\underline{v}}_i^{(o)} \|^2 \leq 2^{n(n-1)/2} |\det V^{(o)}|^2$, and for the row-reduced matrix $\tilde{V}^{(\underline{\lambda})}$

(18) $|\det V^{(\underline{\lambda})}|^2 \leq \prod_{i=1}^{n} \| \tilde{\underline{v}}_i^{(\underline{\lambda})} \|^2 \leq 2^{n(n-1)/2} |\det V^{(\underline{\lambda})}|^2$,

where the bounds in (18) depend indeed not on T. By 1.12 Remark, denoting the Euclidean norm of the smallest vector in the lattice generated by the rows of $V^{(o)}$ (resp. $V^{(\underline{\lambda})}$) by $M_1^{(o)}$ (resp. $M_1^{(\underline{\lambda})}$), we also obtain lower bounds for $\| \tilde{\underline{v}}_i^{(o)} \|^2$ and $\| \tilde{\underline{v}}_i^{(\underline{\lambda})} \|^2$, $(1 \leq i \leq n)$

(19) $\| \tilde{\underline{v}}_i^{(o)} \|^2 \geq (M_1^{(o)})^2 \geq \frac{1}{(2^n\ n\ v^2)^{n-1}} |\det V^{(o)}|^2, \quad (1 \leq i \leq n)$

and considering the transition matrix from $V^{(\underline{\lambda})}$ to $V^{(o)}$

(20) $\| \tilde{\underline{v}}_i^{(\underline{\lambda})} \|^2 \geq (M_1^{(\underline{\lambda})})^2 \geq \frac{1}{\lambda^T (2^n \cdot n\ v^2)^{n-1}} |\det V^{(o)}|^2, (1 \leq i \leq n)$.

Now we can estimate the number $N^{(\underline{\lambda})}$ of vectors $\underline{x} \in \mathbb{Z}^n$ which must be tested for solving (3) after the reduction of the quadratic form $\underline{x}^{tr} U D^2_{\underline{\lambda}} U^{tr} \underline{x}$.

By [2], [3] we have

$N^{(\underline{\lambda})} \leq \frac{1}{2} \prod_{i=1}^{n} (2(\lfloor \| \tilde{\underline{v}}_i^{(\underline{\lambda})} \|^2\ n(K+1)^{2/n} \rfloor)^{1/2} + 1)$.

For T(e.g.B) sufficiently large we conclude from the relations (16), (20) and (18)

$$\leq \frac{1}{2}(2(2^{n-1} \lambda^T\ n\ v^2\ n(K+1)^{2/n})^{\frac{1}{2}} + 1)^{n/2} =: N \text{ . Because of}$$

$T \leq \frac{\log(B^2)}{\log(\lambda)} + 1 - \frac{\log(K^{2/n})}{\log(\lambda)}$ we finally obtain

(21) $N^{(\underline{\lambda})} \leq \frac{1}{2}(2^{\frac{n+1}{2}} B\ n\ v\ \lambda^{1/2} (1 + \frac{1}{K})^{1/n} + 1)^{n/2}$.

Furthermore it is easily seen that the analysis whether a possible solution $\underline{x} \in \mathbb{Z}^n$ actually solves (3) requires at most

(22) $2n^2 + n - 1$

arithmetic operations.

Next the number of arithmetic operations of the reduction algorithm [4] applied to our special problem must be investigated.Using the notations of [4] we have

$$(23)\quad D \leq (n\,\lambda^T v^2)^{\frac{n(n-1)}{2}}$$

and by (20)

$$(24)\quad d_i \geq (\tfrac{3}{4})^{i(i-1)/2} \left(\frac{1}{\lambda^T (2^n\, n\, v^2)^{n-1}} |\det V^{(o)}|^2\right)^i,\quad (1 \leq i \leq n).$$

If B is sufficiently large, the number of arithmetic operations needed by the reduction algorithm of [4] is in our case bounded by

(25) $cn^4 \log(n\, v^2 B^4 \lambda^2)$ with some fixed constant $c \in \mathbb{R}^{>o}$.
This is of a lower order of magnitude than the product of (21) and (22).
In addition the number of arithmetic operations for a change of the quadratic form and for an initialization of the algorithm are negli - gible.

□

On the other hand for $U_i = B$ $(1 \leq i \leq n)$ the enumeration of all $\underline{x} \in \mathbb{Z}^n \setminus \{\underline{0}\}$ subject to (4) of Lemma 1 (the usual procedure of solving (3)) yields $\frac{1}{2}((2\lfloor Bnv \rfloor+1)^n - 1)$ vectors $\underline{x} \in \mathbb{Z}^n$ whose corresponding norms must be calculated. (Here we assume that the sum of the absolute values of the elements of each row of the matrix $((w_j^{(i)})_{1\leq i,j\leq n})^{-1}$ is of size vn.)
Thus the number of arithmetic operations needed by the usual enumeration procedure is greater than

$$\tfrac{1}{2}((2\lfloor Bnv \rfloor + 1)^n - 1)(2n - 1)$$

and in the interesting case of large B, this is about the square of the number of arithmetic operations needed by our procedure.

4. Geometrical interpretation of our method.

The usual way to solve (3) considers all lattice points $\underline{x} \in \mathbb{Z}^n$ subject to (4) of Lemma 1, whereas we take only those for which $|N < \underline{w}, \underline{x} >| \leq K$. The region of $\mathbb{R}^n$ in which they are contained is then suitable covered by ellipsoids as shown in Figure 1 for n = 2.

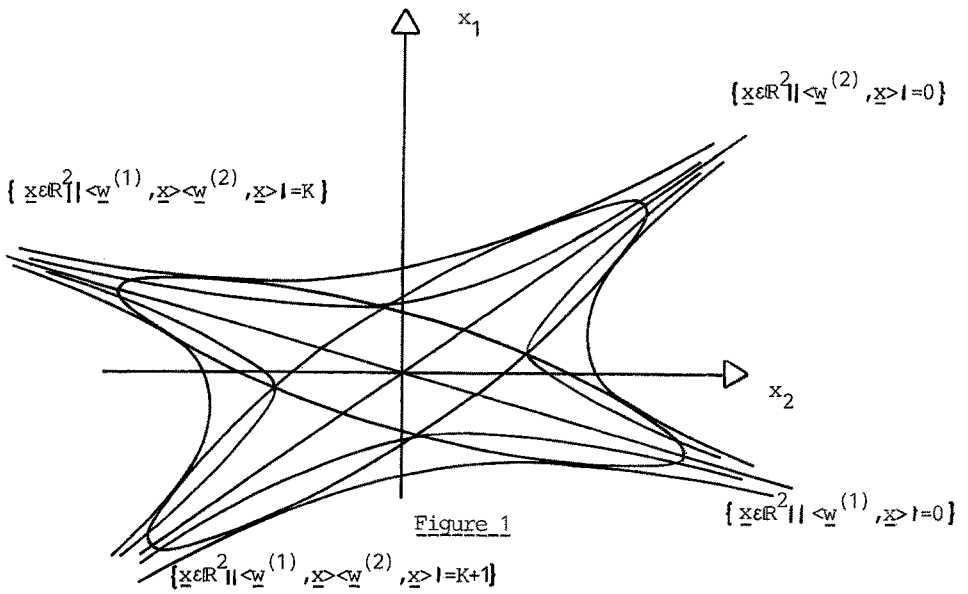

Figure 1

References.

[1] S.I. Borewics und I,R. Safarevič, Zahlentheorie,
Birkhäuser-Verlag, Basel und Stuttgart
1966, pp. 134-141.

[2] U. Dieter, How to Calculate Shortest Vectors in a Lattice,
Math. Comp., v. 29, 131,
(1975), pp. 827-833.

[3] D.E. Knuth, The art of computer programming, vol.2,
Addison-Wesley, sec.ed. (1981), p.95.

[4] A.K. Lenstra, H.W. Lenstr Jr., L. Lovász, Factoring Polynomials
with Rational Coefficients,
Math. Ann. 261 (1982), 515-534.

[5] K. Mahler, Inequalities for Ideal Bases in Algebraic Number Fields,
J. Austral. Math. Soc. 4, (1964), pp. 425-447.

[6] M. Pohst u. H. Zassenhaus, An effective number geometric method
of computing the fundamental units
of an algebraic number field,
Math. Comp., v. 31, 1977, pp. 754-770.

[7] M. Pohst, P. Weiler u. H. Zassenhaus, On effective computation
of fundamental units I,II,
Math. Comp., v. 38, 1982, pp. 275-329.