Computing Greatest Common Divisors and Factorizations in Quadratic Number Fields*

Erich Kaltofen

Rensselaer Polytechnic Institute
Department of Computer Science
Troy, New York 12181

Heinrich Rolletschek

Kent State University
Department of Mathematics
Kent, Ohio 44242

Abstract. In a quadratic number field $\mathbf{Q}(\sqrt{D})$, D a squarefree integer, with class number 1 any algebraic integer can be decomposed uniquely into primes but for only 21 domains Euclidean algorithms are known. It was shown by Cohn [5] that for $D \leq -19$ even remainder sequences with possibly non-decreasing norms cannot determine the GCD of arbitrary inputs. We extend this result by showing that there does not even exist an input in these domains for which the GCD computation becomes possible by allowing non-decreasing norms or remainders whose norms are not as small as possible. We then provide two algorithms for computing the GCD of algebraic integers in quadratic number fields $\mathbf{Q}(\sqrt{D})$. The first applies only to complex quadratic number fields with class number 1, and it is based on a short vector construction in a lattice. Its complexity is $O(S^3)$, where S is the number of bits needed to encode the input. The second allows to compute GCDs of algebraic integers in arbitrary number fields (ideal GCDs if the class number is > 1). It requires only $O(S^2)$ binary steps for fixed D, but works poorly if D is large. Finally, we prove that in any domain the computation of the prime factorization of an algebraic integer can be reduced in polynomial-time to the problem of factoring its norm into rational primes. Our reduction is based on a constructive version of a theorem by A. Thue.

Keywords: quadratic number fields, greatest common divisor, factorization, polynomial-time complexity.

---

* This material is based upon work supported in part by the National Science Foundation under Grant No. DCR-85-04391 and by an IBM Faculty Development Award (first author), and by the Oesterreichische Forschungsgemeinschaft under Grant No. 09/0004 (second author). Appears in *Math. Comput.*, **53**/188, pp. 697-720 (1989). A preliminary version of this article appeared in Springer Lec. Notes Comp. Sci. **204**, 279-288 (1985).

1. Introduction

The fact that the set of complex integers $\{x+\sqrt{-1}\,y \mid x, y \in \mathbf{Z}\}$ forms a unique factorization domain follows from the ability to perform divisions with remainder in this domain. Already C. F. Gauss generalized the complex integers to the domain of algebraic integers

$$O_d = \{ \frac{x+\sqrt{d}\,y}{2} \mid x, y \in \mathbf{Z},\ x \equiv d\ y \bmod 2\},\ \ d = \begin{cases} 4D & \text{if } D \equiv 2,3 \bmod 4 \\ D & \text{if } D \equiv 1 \bmod 4 \end{cases}$$

of a quadratic number field $\mathbf{Q}(\sqrt{D})$, D a squarefree integer, and proved that a Euclidean division (and thus unique factorization) was possible for $d = -11, -8, -7, -4, -3, 5, 8$ and several more positive discriminants d. In general, an abstract integral domain R is Euclidean with respect to a degree function N from the non-zero elements of R into the non-negative integers if for any two elements $\alpha, \beta \in R$, $\beta \neq 0$, either β divides α or there exists a (Euclidean) quotient $\gamma \in R$ and a (Euclidean) remainder $\rho \in R$ such that $\alpha = \beta\gamma + \rho$ and $N(\rho) < N(\beta)$. Once such a Euclidean division is constructible, the GCD of any two elements in R can be determined by repeated division. The Euclidean algorithm (EA) consists of computing for any two elements $\rho_0, \rho_1 \in R$, a sequence of Euclidean divisions

$$\rho_i = \rho_{i-2} - \gamma_{i-1}\rho_{i-1}, \quad i \geq 2,$$

such that $N(\rho_i) < N(\rho_{i-1})$ or $\rho_i = 0$, in which case $\mathrm{GCD}(\rho_0, \rho_1) = \rho_{i-1}$. In the cases of quadratic number fields mentioned before the norm serves as a degree function for the Euclidean algorithm. In this paper we investigate the sequential complexity for GCD computations and prime factorizations in any quadratic number field, including non-Euclidean ones.

We measure the input size in terms of the rational and irrational parts of the inputs. Let

$$R\,\xi = \frac{x}{2}, \ \ I\,\xi = \frac{y\sqrt{|d|}}{2} \ \text{ for } \xi = \frac{x+y\sqrt{d}}{2} \in O_d$$

and let

$$\mathrm{size}(\xi) = \log(|R\,\xi|) + \log(|I\,\xi|),$$

which is the number of bits necessary to write down ξ. With our notation the norm of ξ, $\mathbf{N}\,\xi$, can be represented as follows:

$$\mathbf{N}\,\xi = \xi\,\bar{\xi} = \frac{x+y\sqrt{d}}{2}\ \frac{x-y\sqrt{d}}{2} = \frac{x^2 - d\,y^2}{4} =$$

$$= \left\{ \begin{array}{ll} (R\,\xi)^2 - (I\,\xi)^2 & \text{for } d > 0 \\ (R\,\xi)^2 + (I\,\xi)^2 & \text{for } d < 0 \end{array} \right\} \in \mathbf{Z}.$$

It is easy to show that O_d is a Euclidean domain with respect to the norm for $d = -11, -8, -7, -4, -3$, but for no other $d < 0$, and for $d = 5, 8, 12$ and 13. There are exactly twelve more

discriminants $d > 0$ for which O_d is Euclidean with respect to the norm, namely 17, 21, 24, 28, 29, 33, 37, 41, 44, 57, 73 and 76 (c.f. Chatland/Davenport [4] and Barnes/Swinnerton-Dyer [1]).

O_d is a UFD if and only if its class number $h(d)$ is one. There are exactly four more quadratic number fields with $h(d) = 1$, namely $d = -19, -43, -67$, and -163. That this list is exhaustive was finally established by Stark [29]. It is conjectured that infinitely many real quadratic number fields have class number 1. The list of those $d > 0$ for which O_d is not Euclidean with respect to the norm, but with $h(d) = 1$, begins with 53, 56, 61, 69, 77, 88, 89, 92, 93, $\cdots$.

By definition, the Euclidean algorithm (EA) with respect to the norm applies to ρ_0, ρ_1 exactly if there exists a norm-decreasing remainder sequence $< \rho_1, \ldots, \rho_n >$ with

$$\rho_2 = \rho_0 - \gamma_1 \rho_1,\ \rho_3 = \rho_1 - \gamma_2 \rho_2, \ldots,\ \rho_n = \rho_{n-2} - \gamma_{n-1} \rho_{n-1},\ 0 = \rho_{n-1} - \gamma_n \rho_n, \tag{1.1}$$

where $\gamma_1, \ldots, \gamma_n \in O_d$. The question arises, whether one can introduce another degree function with respect to which O_d with $h(d) = 1$ becomes Euclidean. Even more generally, one may drop the condition that the remainder sequence be decreasing with respect to any degree function and ask if there exists any sequence of the form (1.1), norm-decreasing or not. Before discussing this question, we introduce some more notation.

Generalizing (1.1), we consider a (finite or infinite) sequence $\Sigma = < \rho_0, \rho_1, \rho_2, \ldots >$, where all $\rho_i \in \mathbf{Q}(\sqrt{d})$ (not necessarily $\in O_d$). We define $l = l(\Sigma)$ to be ∞, if Σ is infinite, or $n+1$, if $\Sigma = < \rho_0, \ldots, \rho_{n+1} >$. Following Cooke [6], we call Σ a *division chain*, if

i) For all $i, 1 \le i < l(\Sigma)$, there exists a $\gamma_i \in O_d$ such that $\rho_{i+1} = \rho_{i-1} - \gamma_i \rho_i$,

ii) $\rho_i \neq 0$ for all $i < l$; if $l < \infty$, then $\rho_l = 0$.

The Euclidean algorithm in the most general sense consists of computing a division chain $< \rho_0, \rho_1, \ldots, \rho_{n+1} >$ and returning ρ_n, and every rule for specifying the choice of the remainders defines a *version* of (EA). If the norms of all remainders are minimized, that is, if for all $i < l$, γ_i is chosen such that $\mathbf{N}(\rho_{i+1}) = \mathbf{N}(\rho_{i-1} - \gamma_i \rho_i)$ has the smallest possible value, then we call Σ a *minimal remainder* division chain, the corresponding instance of (EA) a minimal remainder-version of (EA). Notice that we thus relax (1.1) in that we allow remainders with norm larger then the corresponding divisors. We also apply the attribute 'minimal remainder-' to the individual divisions; thus the terms 'minimal remainder-quotient,' 'minimal remainder,' 'minimal remainder-division' have the obvious meaning. In this paper, the elements γ_i will always have the meaning specified in i), whenever a division chain Σ is considered; for sequences Σ', Σ'' etc. we use γ_i', γ_i'' etc., without defining these numbers.

We now turn to the question raised above. First, consider the quadratic domains O_d with $d \le -19$. It is easy to show that the domains O_d, $d = -19, -43, -67$, and -163 are not Euclidean with respect to any degree function (Samuel [24]). Also, there exist ρ_0, ρ_1 such that no finite division chain beginning with ρ_0, ρ_1 exists (Cohn [5]). Thus there does not exist any version of

(EA) which can compute GCD(ρ_0, ρ_1).

Our first group of results extends these facts. We show in section 2 that if $\Sigma = < \rho_0, \rho_1, \ldots >$ is an arbitrary division chain and $\Sigma' = < \rho_0, \rho_1, \rho_2', \rho_3', \ldots >$ a minimal remainder division chain, then $l(\Sigma) \geq l(\Sigma')$. Thus nothing can ever be gained by choosing remainders whose norms are not as small as possible; neither can this reduce the number of divisions needed, nor does there exist even a single input (ρ_0, ρ_1) for which the computation of GCD(ρ_0, ρ_1) succeeds by choosing a non-minimal remainder at some stage, but does not succeed otherwise.

It has previously been shown for certain Euclidean domains that choosing minimal remainders with respect to the standard degree function minimizes the number of divisions. This was done by Lazard for **Z** and for $K[x]$, K a field ([17]), and also for O_{-3} (private communication), and by Rolletschek [22] for O_{-4}. Here the standard degree function is the absolute value for the domain **Z**, the degree in the usual sense for polynomials and the norm for imaginary quadratic number fields. For $d = -4$, that is the Gaussian integers, Caviness/Collins [3] have adopted Lehmer's idea for integer GCD (cf. Knuth [14], §4.5.2), whereas Rolletschek [21], [22] established the equivalent of Lamé's [15] bound on the maximum number of possible divisions necessary.

Our second theorem from section 2 says that in O_d, $d = -19, -43, -67$, and -163, GCD(ρ_0, ρ_1) can be computed by some version of (EA) only if a norm-decreasing sequence of remainders can be achieved. From this the results of Samuel and Cohn follow as an easy corollary.

In the case $d > 0$ the situation is different. Under a generalized Riemann hypothesis, Weinberger [32] shows that every unique factorization domain O_d, $d > 0$, is Euclidean with respect to some degree function, and in Cooke/Weinberger [7] it is shown that a constant bound for the number of divisions can be achieved, namely 5. In fact, these results are shown for the rings of algebraic integers in arbitrary algebraic number fields, provided there infinitely many units. It is not shown, however, how one can efficiently construct these division chains and thus compute GCDs.

There remains the need for efficient algorithms for computing GCDs in those quadratic domains which are not Euclidean with respect to the norm, both real and imaginary. We will describe two such algorithms in this paper. The first, to be presented in section 3, only applies to the imaginary quadratic domains O_d, $d = -19, -43, -67$ or -163. It is based on a short vector construction in a lattice. The number of binary steps needed to compute GCD(ξ, η) is

$$O(S^3), \quad S = \text{size}\,\xi + \text{size}\,\eta.$$

The polynomial bound would still remain valid for variable $d < 0$. In other words, a polynomial upper bound for the complexity can be proved without using the fact shown in [29] that only 4 non-Euclidean imaginary quadratic domains O_d satisfy $h(d) = 1$.

The second algorithm, given in section 4, is much more general, because it applies to all quadratic domains O_d, both real and imaginary, including those with $h(d) > 1$. In the latter case, ideal GCDs are computed in a certain sense, which we will precisely specify later. The algorithm has quadratic complexity for any fixed d. However, it requires some preparatory work (independent of ξ, η), which is quite costly if d is large. It is not clear to us how to accomplish polynomial running time if d is not fixed. Concluding the discussion of GCD-algorithms, we wish to point out that an asymptotically fast algorithm equivalent to Schönhage's [25] integer-half-GCD-algorithm is still not described at this time.

We next turn to factorization into primes. One realizes easily that an algorithm for factoring in O_d can be devised which requires only an additional polynomial cost beyond the factorization of certain rational integers x. The reason is that a rational prime p either remains irreducible in O_d or factors into $p = \pi_1 \pi_2$ $p = \pi_1 \pi_2$ with $\mathbf{N}\,\pi_1 = \mathbf{N}\,\pi_2 = p$. In this paper, we establish that from a factorization of $\mathbf{N}\,\xi$ into primes $p_1 \cdot \cdots p_k$ we can construct for fixed d in deterministic polynomial time the factorization $\xi = \pi_1 \cdot \cdots \pi_n$. We solve this problem by using a constructive version of a theorem by A. Thue for solving the diophantine equation $x^2 - dy^2 = z\,p$ in x, y and small z. This approach follows Shanks [28], Section 71, although our z is about the squareroot of the one obtained there. The main feature of this approach is that it does not require the computation of greatest common divisors in O_d, which is the basis for another standard procedure to solve this problem. That our reduction is deterministic follows from a result by R. Schoof [27] and the fact that d is fixed.

2. Properties of Quadratic Fields

The point of this section is to show that in imaginary quadratic domains O_d, $d \le -19$, one cannot speed up (EA) or increase the set of input for which (EA) works by allowing remainder sequences which are not norm-decreasing or remainders of non-minimal norm. First we recall some well-known facts about quadratic fields, the proofs of which can be found, e. g. in Hasse's text [9], §16. Let

$$\omega_d = \begin{cases} \dfrac{1+\sqrt{d}}{2} & \text{for } d \equiv 1 \bmod 4, \\ \dfrac{\sqrt{d}}{2} & \text{for } d \equiv 0 \bmod 4. \end{cases}$$

Then the set $\{1, \omega_d\}$ forms an integral basis for O_d, that is, every element of O_d can be represented in the form $a \cdot 1 + b \cdot \omega_d$, with $a, b \in \mathbf{Z}$.

A *unit* $\varepsilon \in O_d$ is an element such that $1/\varepsilon \in O_d$. A necessary and sufficient condition for ε to be a unit is that $|\mathbf{N}\,\varepsilon| = 1$. For $d < 0$ the multiplicative group of units is generated by $\{\sqrt{-1}\}$ for $d = -4$, $\{(1+\sqrt{-3})/2\}$ for $d = -3$ and $\{-1\}$ in all other cases. For $d > 0$ there always exists a *fundamental unit* ε_1, $R\,\varepsilon_1 > 0$, $I\,\varepsilon_1 > 0$, such that the unit group is generated by $\{-1, \varepsilon_1\}$. Two elements $\xi_1, \xi_2 \in O_d$ are *associates*, $\xi_1 \sim \xi_2$, if there exists a unit ε such that $\xi_1 = \varepsilon\xi_2$. For $d > 0$ any ξ_1 has, according to [9], an associate ξ_2 with

$$|R\,\xi_2|, |I\,\xi_2| < \frac{|\mathbf{N}\,\xi_1| + \varepsilon_1}{2}.$$

However, we will need a better estimate in §4. By ϕ we denote the argument-function, which is defined for all complex numbers $\xi \neq 0$ by $\xi = |\xi| \cdot (\cos(\phi(\xi)) + i \sin(\phi(\xi)))$. We also need the analogous definition given in [9], p. 288, for elements of real quadratic fields: for $d > 0$, $\phi(\xi)$ is defined by $\xi = \mathrm{sign}(\xi)\sqrt{|\mathbf{N}\xi|}e^{\phi(\xi)}$; then $\phi(\bar{\xi}) = -\phi(\xi)$. Dividing ξ_1 by a power ε_1^k of ε_1 such that $\phi(\varepsilon_1^k)$ is as close to $\phi(\xi_1)$ as possible, we can also find an associate ξ_2 of ξ_1 such that $|\phi(\xi_2)| \le \phi(\varepsilon_1)/2$. By adding and subtracting the equalities

$$\xi_2 = \mathrm{sign}(\xi_2)\sqrt{|\mathbf{N}\xi_2|}e^{\phi(\xi_2)},$$

$$\bar{\xi}_2 = \mathrm{sign}(\bar{\xi}_2)\sqrt{|\mathbf{N}\xi_2|}e^{-\phi(\xi_2)}$$

we get $2|R\xi_2|, 2|I\xi_2| \le \sqrt{|\mathbf{N}\xi_2|} \cdot (e^{\phi(\xi_2)} + e^{-\phi(\xi_2)})$, hence

$$|R\xi_2|, |I\xi_2| \le \sqrt{|\mathbf{N}\xi_2|}e^{|\phi(\xi_2)|} \le \sqrt{|\mathbf{N}\xi_2|}\sqrt{\varepsilon_1}.$$

We now discuss how rational primes p split in O_d, $h(d) = 1$. If $p \mid d$ then $p \sim \pi^2$ for some prime $\pi \in O_d$. If $p \ge 3$ and the Legendre symbol $(d/p) = (D/p) = +1$ then there exists $l \in \mathbf{Z}$ such that $l^2 \equiv D \bmod p$. Therefore $p \mid (l+\sqrt{D})(l-\sqrt{D})$ and thus $p \sim \pi\,\bar{\pi}$ with $\pi = \mathrm{GCD}(p,$

$l + \sqrt{D}$). If $p = 2$ and $d = D \equiv 1 \bmod 8$ then

$$2 \mid (1 - \frac{1+\sqrt{D}}{2})\,(1 - \frac{1-\sqrt{D}}{2})$$

and thus $2 \sim \pi\,\bar{\pi}$ with $\pi = \mathrm{GCD}(2, 1 - (1 + \sqrt{D})/2)$. In all other cases p is a prime in O_d.

We now come to the main results of this section, which apply to the domains O_d, $d = -19$, -43, -67, -163. Although one would usually expect that minimal remainder-versions require the smallest number of divisions among all versions of (EA), one might suspect that there are exceptional instances, where other versions terminate faster. The following theorem shows, however, that this is not the case in those imaginary quadratic fields under consideration.

Theorem 2.1: Let $\mathbf{Q}(\sqrt{d})$ be an imaginary quadratic number field whose corresponding number ring O_d is a UFD without a Euclidean algorithm with respect to the norm, that is, d has one of the values –19, –43, –67, –163. Let $\Sigma = < \rho_0, \rho_1, \rho_2, \ldots >$ be an arbitrary division chain, and let $\Sigma' = < \rho_0, \rho_1, \rho_2', \rho_3', \ldots >$ be a minimal-remainder division chain beginning with the same two elements ρ_0, ρ_1. Then $l(\Sigma) \geq l(\Sigma')$.

Proof: We apply a technique developed in Lazard [17] and in Rolletschek [22]. In what follows, we call a division chain $\Sigma = < \rho_0, \rho_1, \ldots >$ a counterexample, if there exists a sequence $\Sigma' = < \rho_0, \rho_1, \rho_2', \ldots >$ such that the assertion of the theorem does not hold. Since the theorem is trivially true for $l(\Sigma) = \infty$, we may apply induction on $l(\Sigma)$. Thus assume that $\Sigma = < \rho_0, \rho_1, \ldots, \rho_{n+1} >$ is a counterexample, but that the theorem is true for every shorter sequence in place of Σ. Let Σ' be a minimal remainder-division chain for ρ_0, ρ_1 with $l(\Sigma) < l(\Sigma')$. Then $n \geq 2$ and $\rho_2 \neq \rho_2'$, otherwise $< \rho_1, \rho_2, \ldots, \rho_{n+1} >$ would be another counterexample, contrary to the induction hypothesis. Without loss of generality we may make the following three assumptions:

i) All remainders in Σ except possibly ρ_2 are minimal. For otherwise we could have considered a minimal remainder-division chain $< \rho_1, \rho_2, \rho_3'', \ldots >$ whose length k, by induction hypothesis, would have to be no more than n; but then $< \rho_0, \rho_1, \rho_2, \rho_3'', \ldots, \rho_{k+1}'' >$ would be another counterexample which would also satisfy i).

ii) $\gamma_1' = 0$; otherwise we could replace ρ_0 by $\rho_0 - \gamma_1'\rho_1$, γ_1' by 0 to get a counterexample satisfying ii).

iii) $\rho_1 = 1$; otherwise we could divide $\rho_0, \ldots, \rho_{n+1}$ by ρ_1 to obtain a counterexample satisfying iii). It is here that the consideration of non-integral values of $\rho_0, \ldots$ makes the proof more convenient.

It follows from assumption ii) that 0 is one of the algebraic integers in O_d closest to $\rho_0 = \rho_2$. Hence ρ_0 must lie in a region $R_1 = \{\alpha \mid \alpha \in \mathbf{Q}(\sqrt{d}), |\alpha| \leq |\alpha - \gamma| \text{ for all } \gamma \in O_d\}$, which is shown in fig. 1.

Place figure 1 here or below

Since $\rho_2 \neq \rho_2'$, $\gamma_1 \neq 0$. We consider the various possible values of γ_1. In several cases we have to consider the value of $\rho_1/\rho_2 = 1/\rho_2$, and we denote this value by δ.

a) $\gamma_1 = 1$. Then ρ_2 lies in the region R_2 which is constructed by shifting R_1 to the left by 1; R_2 is bounded by the straight lines $-3/2 + yi$ and $-1/2 + yi$ ($y \in \mathbf{R}$), and 4 additional straight lines. Hence δ certainly lies within the region R_2' shown in fig. 2;

Place figure 2 here or below

here the circles C_1, C_2 are the sets of inverses of all points of the form $-1/2 + yi$ and of the form $-3/2 + yi$ respectively.

Now recall the definition of the element ω_d at the beginning of this section. We need the fact that $I\,\omega_d > 2$. δ has a distance ≤ 1 from -1, but a distance > 1 from all lattice points outside the real axis. Hence γ_2, which is a minimal remainder-quotient of ρ_1 and ρ_2 by assumption i), and which is therefore one of the elements of O_d closest to δ, can only have one of the values 0, -1 or -2. Correspondingly, we have to consider three subcases.

a1) $\gamma_2 = 0$. Then the sequence Σ starts with $< \rho_0,\ 1,\ \rho_0 - 1,\ 1,\ \rho_0,\ \rho_5, \ldots >$, since the minimal remainder of $\rho_0 - 1$ and 1 equals the minimal remainder of ρ_0 and 1. (More precisely, we may assume without loss of generality that $\rho_4 = \rho_0$ by the same argument as in the justification of assumption i), although there may be several minimal remainders of ρ_0 and 1.) We now consider the division chain $\Sigma'' = < \rho_0,\ 1,\ \rho_0,\ \rho_5, \ldots,\ \rho_{n+1} >$. Σ'' is actually a minimal remainder-division chain, though not necessarily identical with $\Sigma' = < \rho_0,\ 1,\ \rho_0,\ \rho_3', \ldots >$. We can apply the induction hypothesis to the sequences formed from Σ'' and Σ' by omitting their first elements; since Σ' is also a minimal remainder-division chain, it follows that $l(\Sigma') \leq l(\Sigma'') = n - 1$, contradicting the assumption that $n + 1 = l(\Sigma) < l(\Sigma')$. This concludes the proof of the theorem for this case.

a2) $\gamma_2 = -1$. In this case, Σ starts with $< \rho_0,\ 1,\ \rho_0 - 1,\ \rho_0,\ \rho_4, \ldots >$. Again we construct an division chain Σ'' shorter than Σ: $\Sigma'' = < \rho_0, 1,\ \rho_0, -\rho_4,\ \rho_5, \ldots,\ (-1)^n \rho_{n+1} >$, this time using the fact that $1 \bmod \rho_0 = -((\rho_0 - 1) \bmod \rho_0)$. Then we can show as in a1) that the given minimal remainder-division chain Σ' satisfies $l(\Sigma') \leq l(\Sigma'') < l(\Sigma)$, a contradiction.

a3) $\rho_2 = -2$. The sequence Σ starts with

$$< \rho_0, 1,\ \rho_0 - 1,\ 2\rho_0 - 1,\ \rho_0(1 - 2\gamma_3) + (-1 + \gamma_3),\ \rho_5,\ \rho_6, \ldots > .$$

We put

$$\Sigma'' = < \rho_0, 1,\ \rho_0, 1 - 2\rho_0,\ \rho_0(-1 + 2\gamma_3) + (1 - \gamma_3), -\rho_5, -\rho_6, \ldots > ,$$

choosing $\gamma_3'' = \gamma_3 - 1$. In this case, $l(\Sigma') \leq l(\Sigma'') = l(\Sigma)$, a contradiction.

b) $\gamma_1 = -1$. This case parallels a) completely, only some signs have to be changed.

c) $\gamma_1 = 2$. As in case a), we can determine the set of all possible values of δ; it is the region R_3' shown in fig. 3.

Place figure 3 here or below

It follows that γ_2 can only be 0 or –1. The case $\gamma_2 = 0$ is treated as in case a1). Assume $\gamma_2 = -1$. Σ now has the form

$$< \rho_0, 1, \rho_0 - 2, \rho_0 - 1, \rho_0(1 - \gamma_3) + (-2 + \gamma_3), \rho_5, \ldots > ,$$

and the sequence Σ' we construct has the form

$$< \rho_0, 1, \rho_0, 1 - \rho_0, \rho_0(-1 + \gamma_3) + (2 - \gamma_3), -\rho_5, \ldots > ,$$

where we choose $_3\gamma' = -2 + \gamma_3$. Then $l(\Sigma') = l(\Sigma)$, and the rest of the proof parallels previous cases.

d) $\gamma_1 = -2$. This case is analogous to c).

e) γ_1 is real, $|\gamma_1| \geq 3$. Then $|\rho_2| = |\rho_0 - \gamma_1| > 2$, hence $|\delta| < 1/2$. Then $\gamma_2 = 0$, and the assertion of the theorem follows as in a1).

f) γ_1 is not real: $\gamma_1 = a + bi$ with $b \neq 0$. We show in all cases $|I(\rho_2)| = |I(\rho_0 - \gamma_1)| > 1$. Indeed, the minimal absolute value of $I(\rho_0 - \gamma_1)$ occurs with $d = -19$, $\gamma_1 = \omega_d$ and $\rho_0 = \zeta$ as pictured in fig. 1, and in this case $|I(\rho_2)| > 1$. Fig. 4 shows the region R_4', which is the set of all inverses of complex numbers α with $|I(\alpha)| > 1$, and which contains the set of all possible values of δ.

Place figure 4 here or below

Again it follows that the element $\gamma_2 \in O_d$ closest to δ can only be 0, so $\Sigma = < \rho_0, 1, \rho_0 - \gamma_1, \rho_4, \ldots >$. But now $\Sigma''' = < \rho_0, 1, \rho_4, \ldots >$ is a shorter division chain than Σ, leading to the same contradiction as in the previous cases.

The cases a)-f) are exhaustive, completing the proof. □

The condition $d \leq -19$ in the previous theorem is used in the estimates in the cases a) and f) in the above proof. Remarkably, the statement of the theorem (without the assumption of O_d being non-Euclidean) fails for $d = -11$, but it remains valid for all other Euclidean imaginary number fields, as was recently shown in [23].

In algorithm-theoretic terms, theorem 2.1 can be formulated as follows:

Corollary 2.1: Let d be as in the theorem. Then

i) All minimal remainder-versions of (EA), applied to ρ_0, ρ_1 require the same number of divisions, and no other version requires fewer.

ii) If no minimal remainder-version of (EA) allows the computation of GCD(ρ_0, ρ_1), then no division chain terminates with GCD(ρ_0, ρ_1). □

While theorem 2.1 and its corollary deal with divisions where the remainder has non-minimal norm, the next theorem shows that nothing is gained by allowing divisions where the remainder has norm greater than the divisor, even if that remainder is minimal. In other words, if a minimal remainder-version of (EA), applied to ρ_0, ρ_1, leads to a division where no remainder with norm smaller than the divisor exists, then every version of (EA) fails for this input.

Theorem 2.2: For any $d \leq -19$, $h(d) = 1$, and for all ρ_0, $\rho_1 \in O_d$ there exists a sequence γ_1 ,…, $\gamma_n \in O_d$ satisfying (1.1) only if a norm-decreasing sequence with this property exists, that is, if the common version of (EA) applies to ρ_0, ρ_1, where the remainder of each division has smaller norm than the divisor.

Proof: Consider a minimal remainder-division chain $\Sigma = < \rho_0, \rho_1, \rho_2, \ldots >$, and assume that $< \rho_1, \rho_2, \ldots >$ is not norm-decreasing. We will show that there exists an infinite minimal remainder-division chain Σ' starting with ρ_0, ρ_1. Similarly as in the proof of theorem 2.1. a number of assumptions can be made without loss of generality:

i) $|\mathbf{N}\rho_2| \geq |\mathbf{N}\rho_1|$;
ii) $\rho_1 = 1$;
iii) $\gamma_1 = 0$;
iv) $R\rho_0 \geq 0$, $I\rho_0 \geq 0$.

The justification of i) is immediate. For ii) and iii) see the analogous assumptions in the proof of theorem 2.1. Finally, iv) can be justified by symmetry; in the following proof only some signs and limits for $\phi(\rho_0)$ would have to be changed if $R\rho_0$ and/or $I\rho_0$ is < 0. We note in passing that an assumption analogous to iv) could also have been made in the proof of theorem 2.1., but its justification would have required a few lines, and the rest of the proof would not have been simplified.

We have $\rho_2 = \rho_0$ by iii), $\mathbf{N}(\rho_0) = \mathbf{N}(\rho_2) \geq 1$ by i) and ii). By iii), 0 is one of the algebraic integers in O_d closest to $\rho_0 = \rho_2$. Together these facts imply that ρ_0 lies in the region shown in fig. 5.

Place figure 5 here or below

It follows now that $\pi/3 \leq \phi(\rho_0) \leq \pi/2$, hence $\delta = 1/\rho_0$ satisfies $-\pi/2 \leq \phi(\delta) \leq -\pi/3$; also, $\mathbf{N}(\delta) \leq 1$. Then one of the algebraic integers $\gamma_2 \in O_d$ closest to δ is 0. Choosing $\gamma_2 = 0$, we get $\rho_3 = 1$. Hence the following is a minimal remainder-division chain:

$$\Sigma' = < \rho_0, 1, \rho_0, 1, \rho_0, \ldots > .$$

This sequence is infinite, as desired. By theorem 2.1. every division chain starting with ρ_0, ρ_1 is infinite, so no version of (EA) can terminate.

Corollary 2.2: O_d, $d \leq -19$, $h(d) = 1$, is not a Euclidean domain for any choice of degree function.

Proof: Recall that for $d \leq -19$, O_d is not Euclidean with respect to the norm. This means, by definition, that there exist elements ρ_0, $\rho_1 \in O_d$ such that every remainder of ρ_0 and ρ_1 has larger norm than ρ_1. Thus if Σ is a minimal remainder-division chain beginning with ρ_0, ρ_1, then $< \rho_1, \rho_2, \ldots >$ is not norm-decreasing, and by theorem 2.2 no finite division chain beginning with ρ_0, ρ_1 exists. The assertion follows. □

3. GCD Computation by Lattice Reduction

As we have seen in §2 not all quadratic fields with unique factorization allow a Euclidean algorithm, e.g. $\mathbf{Q}(\sqrt{-19})$ and $\mathbf{Q}(\sqrt{53})$. Therefore a different GCD procedure is required for these domains. Our first algorithm for imaginary quadratic fields is based on computing short integral lattice vectors and is interesting for two reasons, even though the algorithm in §4 has lower asymptotic complexity. First, it does not require any field-dependent preconditioning, and second its running time is polynomial independently of the fact that $|d|$ is known to be bounded. The main idea of the algorithm is to solve $\lambda\xi + \mu\eta = 0$ such that $\mathbf{N}\mu$ is small. The following lemma will be useful.

Lemma 3.1: Let $\xi, \eta \in O_d$, $d \le -19$, $h(d) = 1$, $\xi\eta \ne 0$, $\delta = \mathrm{GCD}(\xi, \eta)$, $\xi^* = \xi/\delta$, $\eta^* = \eta/\delta$. Assume that $\lambda\xi + \mu\eta = 0$, $\lambda, \mu \in O_d$. Then $\eta^* \mid \lambda$, $\xi^* \mid \mu$. Furthermore, if λ is not an associate of η^* then $\mathbf{N}\mu \ge 4\,\mathbf{N}\xi^*$.

Proof: Since $\lambda\xi^* = -\mu\eta^*$ and $\mathrm{GCD}(\xi^*, \eta^*) = 1$ all prime factors of η^* must occur in λ and similarly for ξ^* and μ. Since 2 and 3 are primes in any O_d in question, $\mathbf{N}(\mu/\xi^*) \ge 4$. □

We consider the case $d \equiv 1 \bmod 4$ only, since only this one occurs. Let $\lambda = (l_1 - l_2/2) + \sqrt{d}\, l_2/2$, $\xi = (x_1 - x_2/2) + \sqrt{d}\, x_2/2$, $\mu = (m_1 - m_2/2) + \sqrt{d}\, m_2/2$, $\eta = (y_1 - y_2/2) + \sqrt{d}\, y_2/2$ with l_1, l_2, m_1, m_2, x_1, x_2, y_1, $y_2 \in \mathbf{Z}$. Then

$$\begin{aligned} 4R(\lambda\xi + \mu\eta) &= (4x_1 - 2x_2)l_1 + (x_2 - 2x_1 + dx_2)l_2 + (4y_1 - 2y_2)m_1 + (y_2 - 2y_1 + dy_2)m_2 \\ \frac{2I(\lambda\xi + \mu\eta)}{\sqrt{d}} &= x_2 l_1 + (x_1 - x_2)l_2 + y_2 m_1 + (y_1 - y_2)m_2. \end{aligned} \tag{3.1}$$

We want to find integers l_1, l_2, m_1 and m_2 such that the right sides of the equations (3.1) are 0, meaning that for the corresponding λ and μ, $\lambda\xi + \mu\eta = 0$. Simultaneously we want to keep $\mathbf{N}\mu$ small because the smallest such μ is an associate of ξ^* by the previous lemma. This leads to the problem of finding a short vector in an integer lattice. The next theorem shows that only associates of ξ^* can correspond to short vectors in a particular lattice and thus finding a short vector in that lattice with a reduction algorithm actually gives ξ^* and hence $\delta = \xi/\xi^*$.

Theorem 3.1: Let O_d a UFD with $d \le -19$, and let $\xi = (x_1 - x_2/2) + \sqrt{d}\, x_2/2$ and $\eta = (y_1 - y_2/2) + \sqrt{d}\, y_2/2$ be two nonzero elements of O_d. Let c and $\hat{d}$ be integers such that $c \ge \sqrt{12\,\mathbf{N}\xi}$, $|\sqrt{|d|} - \hat{d}| \le 1/2$. Furthermore, let L^* be the 4-dimensional integer lattice spanned by the columns of the matrix

$$L = \begin{bmatrix} c(4x_1 - 2x_2) & c(x_2 - 2x_1 + dx_2) & c(4y_1 - 2y_2) & c(y_2 - 2y_1 + dy_2) \\ cx_2 & c(x_1 - x_2) & cy_2 & c(y_1 - y_2) \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & \hat{d} \end{bmatrix}.$$

Then if $v^* = L \times [l_1, l_2, m_1, m_2]^T \ne \mathbf{0}$ is a vector of shortest Euclidean length in the lattice L^*, then $\mu = (m_1 - m_2/2) + \sqrt{d}\, m_2/2$ must be an associate of $\xi^* = \xi/\mathrm{GCD}(\xi, \eta)$. Furthermore, if $v \in L^* \setminus$

$\{\mathbf{0}\}$ does not have shortest Euclidean length then $\|v\|^2 > 12/5\,\|v^*\|^2$.

Proof: Since

$$\det\begin{bmatrix} 4x_1 - 2x_2 & x_2 - 2x_1 + dx_2 \\ x_2 & x_1 - x_2 \end{bmatrix} = (2x_1 + x_2)^2 - dx_2^2 > 0, \quad x_1 x_2 \neq 0,$$

L is of full rank. With λ and μ as above we get from (3.1) that

$$L \times \begin{vmatrix} l_1 \\ l_2 \\ m_1 \\ m_2 \end{vmatrix} = \begin{vmatrix} 4cR(\lambda\xi + \mu\eta) \\ 2cI(\lambda\xi + \mu\eta)/\sqrt{d} \\ 2m_1 - m_2 \\ \hat{d}\,m_2 \end{vmatrix} = v = \begin{vmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{vmatrix} \in \mathbf{Z}^4.$$

We now estimate the Euclidean length $\|v\|$ of $v \neq \mathbf{0}$. If $\lambda\xi + \mu\eta \neq 0$ then $\|v\|^2 \geq c^2 \geq 12\,\mathbf{N}\xi$. In case $\lambda\xi + \mu\eta = 0$, $v_1 = v_2 = 0$, and we get from $\|d| - \hat{d}^2| \leq \hat{d}$, $\hat{d}/|d| < 1/4$, $d \leq -19$, and $m_2^2 \leq 4\mathbf{N}\mu/|d|$:

$$\begin{aligned} |\|v\|^2 - 4\mathbf{N}\mu| &= |(2m_1 - m_2)^2 + \hat{d}^2 m_2^2 - ((2m_1 - m_2)^2 - dm_2^2)| \\ &= |\hat{d}^2 - |d\|\, m_2^2 \leq \hat{d}\,m_2^2 \leq \frac{4\hat{d}\mathbf{N}\mu}{|d|} < \mathbf{N}\mu. \end{aligned} \tag{3.2}$$

From (3.2) we conclude that

$$3\mathbf{N}\mu < \|v\|^2 < 5\mathbf{N}\mu \quad \text{for } \lambda\xi + \mu\eta = 0. \tag{3.3}$$

Therefore, under the assumption that μ is an associate of ξ^* and $\lambda\xi + \mu\eta = 0$, the corresponding vector v^* satisfies $\|v^*\|^2 < 5\mathbf{N}\mu = 5\mathbf{N}\xi^*$. Otherwise lemma 3.1 states that $\mathbf{N}\mu \geq 4\mathbf{N}\xi^*$ and thus we get from (3.3)

$$\|v\|^2 \geq \min(12\mathbf{N}\xi, 3\mathbf{N}\mu) \geq 12\mathbf{N}\xi^*.$$

Therefore

$$\|v\|^2 \geq 12\mathbf{N}\xi^* > \frac{12}{5}\,\|v^*\|^2$$

which proves the theorem. □

The algorithm for computing the GCD of ξ and η is now easy. One computes a vector in the lattice L^* whose length is within a factor $C^{3/2}$ of the length of the shortest vector, where C is a constant with $C > 4/3$ and $C^3 < 12/5$, e.g., $C = 83/62$. There exist several versions of the basis reduction algorithm by A. K. Lenstra et al [18],. cf [12].. Already the original algorithm [18] can find such a vector. Since the dimension is fixed all these algorithms take $O(\text{size}^3\xi\eta)$ binary steps to compute such a vector (slightly less if fast multiplication is used). But by theorem 3.1 such a vector must be a shortest vector whose entries determine an associate of ξ^*. We then obtain $\delta = \text{GCD}(\xi, \eta) = \xi/\xi^*$.

4. Greatest Common Divisor Computation

We present a GCD procedure that works for all quadratic domains. It turns out that one always can divide an appropriate multiple $l\,\xi$, $l \in \mathbf{Z}$, of the dividend ξ by η and accomplish a remainder of norm smaller than $|\mathbf{N}\,\eta|$. Moreover, the size of l only depends on d and not on ξ or η. This fact was shown for arbitrary number fields by Hurwitz, see [11], p. 237. The following lemma provides the specific bound for l for quadratic number fields O_d. It is not restricted to the case $h(d) = 1$.

Lemma 4.1: Let $c \in \mathbf{R}$, $1/2 \le c \le 1$, O_d the ring of integers in a quadratic number field. Then for all $\xi, \eta \in O_d$ there exists an l such that

$$l \in \mathbf{Z} \text{ with } 1 \le l \le \left\lfloor \frac{\sqrt{|d|}}{2c} \right\rfloor,$$

and there exists a $\gamma \in O_d$ with

$$|R(l\xi/\eta - \gamma)| \le \frac{1}{2}, \ |I(l\xi/\eta - \gamma)| < c. \tag{4.1}$$

Furthermore, if $d > 0$ then $|\mathbf{N}(l\xi - \gamma\eta)| \le c^2|\mathbf{N}\eta|$; if $d < 0$ then $|\mathbf{N}(l\xi - \gamma\eta)| < (c^2 + 1/4)\,|\mathbf{N}\eta|$.

Proof: This follows from the theory of approximation of real by rational numbers and of continued fractions. We apply theorem 171 in Hardy/Wright [8]: if p_n/q_n and p_{n+1}/q_{n+1} are the n-th and $n+1$-th continued fraction approximation of a real number x, then

$$\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}} \ (< \frac{1}{q_n^2}).$$

We apply this theorem to $x = \dfrac{I(\xi/\eta)}{I(\omega_d)} = \dfrac{I(\xi/\eta)}{\sqrt{|d|}/2}$, and a continued fraction approximation p_n/q_n of x such that $q_n \le \dfrac{\sqrt{|d|}}{2c}$ and either $p_n/q_n = x$ or $q_{n+1} > \sqrt{|d|}/2c$. Then

$$\left| \frac{I(\xi/\eta)}{\sqrt{|d|}/2} - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}} < \frac{c}{q_n \dfrac{\sqrt{|d|}}{2}},$$

hence

$$\left| I(q_n\xi/\eta - p_n \frac{\sqrt{|d|}}{2}) \right| = \left| q_n I(\xi/\eta) - p_n \frac{\sqrt{|d|}}{2} \right| < c.$$

Then (4.1) is satisfied for $l = q_n$, $\gamma = p_n\omega_d + \lfloor R(q_n\xi/\eta - p_n\omega_d + 1/2) \rfloor$.

Now let $\tau = l\xi/\eta - \gamma$. For $d > 0$, $|\mathbf{N}\tau| = |R(\tau)^2 - I(\tau)^2| \le \max(R(\tau)^2, I(\tau)^2) = c^2$. For $d < 0$, $|\mathbf{N}\tau| = R(\tau)^2 + I(\tau)^2 \le c^2 + 1/4$. Then the second statement of the lemma follows by multiplying

these inequalities by $|\mathbf{N}\eta|$. □

This lemma suggests the following algorithm for computing GCD's in arbitrary quadratic domains. We formulate it first for those d with unique factorization. At the end of this section we will discuss how to adapt it to apply to the case $h(d) > 1$.

Preconditioned GCD in O_d

Input: $\xi, \eta \in O_d$, O_d a UFD.

Preconditioning: Given are the prime factors of $l = 2, \ldots, m = \lfloor \sqrt{|d|} \rfloor$ in O_d,

$$l = \pi_1^{e_{l,1}} \cdot \cdot \cdot \pi_k^{e_{l,k}}, \ 2 \le l \le m, \ e_{l,j} \ge 0.$$

If $d > 0$ we furthermore are given a fundamental unit ε_1.

Output: $\delta = \mathrm{GCD}(\xi, \eta)$.

Step 1: $\Pi \leftarrow \varnothing$; $\rho_0 \leftarrow \xi$; $\rho_1 \leftarrow \eta$.
FOR $i \leftarrow 1, 2, \cdot \cdot$ WHILE $\rho_i \neq 0$ DO step 2.

Step 2: Here we carry out a ''pseudo-remainder'' step, that is we compute $l_{i+1} \in \mathbf{Z}$, $1 \le l_{i+1} \le m$, and $\gamma_{i+1} \in O_d$ such that

$$|\mathbf{N}(l_{i+1}\rho_{i-1} - \gamma_{i+1}\rho_i)| < \frac{1}{2}\,|\mathbf{N}\rho_i|.$$

Set $(x_i + y_i\sqrt{d})/2 \leftarrow \rho_{i-1}/\rho_i$, where $x_i, y_i \in \mathbf{Q}$.
Determine l_{i+1} satisfying lemma 3.1 for $c = 1/2$ as follows. Compute the n-th convergent p_n/q_n of the continued fraction approximation for y_i such that $q_n \le m$ and either $p_n/q_n = y_i$ or $q_{n+1} > m$.
Set $l_{i+1} \leftarrow q_n$; $y_{i+1} \leftarrow p_n$. At this point $|l_{i+1}y_i - y_{i+1}| < 1/\sqrt{d}$.
Compute $x_{i+1} \in \mathbf{Z}$ such that

$$\left| \frac{x_i - (dy_{i+1} \bmod 2)}{2} - x_{i+1} \right| \le \frac{1}{2}.$$

Set $\gamma_{i+1} \leftarrow x_{i+1} + (dy_{i+1} \bmod 2 + y_{i+1}\sqrt{d})/2$.
Set $\rho_{i+1} \leftarrow l_{i+1}\rho_{i-1} - \gamma_{i+1}\rho_i$; $\Pi \leftarrow \Pi \cup \{\pi_j \mid 1 \le j \le k, \ e_{l,j} \ge 1\}$.
At this point $|\mathbf{N}\rho_{i+1}| < |\mathbf{N}\rho_i|/2$.
IF $d > 0$ and $|\mathbf{N}\rho_{i+1}| \neq 1$ THEN

> We adjust ρ_{i+1} such that $R\,\rho_{i+1}$ and $I\,\rho_{i+1}$ do not become too large, as follows:
> Compute $k \ge 0$ such that $|\phi(\varepsilon_1^k) - \phi(\rho_{i+1})| < \phi(\varepsilon_1)/2$.
> Set $\rho_{i+1} \leftarrow \rho_{i+1}/\varepsilon_1^k$. At this point $|R\,\rho_{i+1}|$ and $|I\,\rho_{i+1}|$ are both $< \sqrt{|\mathbf{N}\rho_{i+1}|}\sqrt{\varepsilon_1}$. (Refer to §2 for an explanation of these facts.)

Step 3: Remove extraneous factors from ρ_{i-1} introduced by the l_i.
Set $\beta \leftarrow 1$. FOR $\pi \in \Pi$ DO

Compute the maximal $e,\ f$ such that $\pi^e \mid \xi,\ \pi^f \mid \eta$.
Set $\beta \leftarrow \beta \times \pi^{\min(e,f)}$; $\rho_{i-1} \leftarrow \rho_{i-1}/\pi^{\min(e,f)}$.
WHILE $\pi \mid \rho_{i-1}$ DO $\rho_{i-1} \leftarrow \rho_{i-1}/\pi$.

RETURN $\delta \leftarrow \beta\, \rho_{i-1}$. □

If one also needs the extended Euclidean scheme $\sigma\xi + \tau\eta = \delta$ it suffices to set $\sigma = (s_1 + \sqrt{d}\, s_2)/2,\ \tau = (t_1 + \sqrt{d}\, t_2)/2,\ s_1 - ds_2 = 2s_3,\ t_1 - dt_2 = 2t_3$, and solve the four resulting linear equations with integer coefficients in the integers $s_1,\ s_2,\ s_3,\ t_1,\ t_2,\ t_3$ by some integer linear system solver, see R. Kannan and A. Bachem [13].

It is easy to show that for fixed d this algorithm has time complexity $O(S^3)$, where S is the size of the input, defined in §1. Some additional effort is needed to show that the complexity can be improved to $O(S^2)$, and lastly to determine the complexity in dependence of d.

Theorem 4.1: After input independent preconditioning, the GCD of $\xi,\ \eta \in O_d$, O_d a UFD, can be computed in

$$O(S\ \sqrt{|d|}(\log d)\ (S + \sqrt{|d|}\log d)), \quad S = \text{size}\,\xi + \text{size}\,\eta,$$

binary steps using classical integer arithmetic procedures.

Proof: This can be shown similarly as for the Euclidean algorithm applied to rational integers (see Knuth [14], Exercise 30 of §4.5.2), but the details are somewhat more involved. We have to analyze the preconditioned algorithm, and also to apply some minor modifications. In what follows, we mean by 'constant' some quantity that is independent of ξ and η, though it may still depend on d. First we have to consider the number of iterations of step 2. It follows from $|\mathbf{N}\rho_{i+1}| < |\mathbf{N}\rho_i|/2$ that this number is bounded by $\log|\mathbf{N}\rho_1| + 1$, which is $O(S)$.

Second, we have to establish an upper bound for the size of the remainders $\rho_2,\ \rho_3$ etc. Because of the adjustment of ρ_{i+1} at the end of each iteration of step 2, applied for $d > 0$, there exists a constant C_1 such that both $|R(\rho_{i+1})|$ and $|I(\rho_{i+1})| \le C_1\sqrt{|\mathbf{N}\rho_{i+1}|} < C_1\sqrt{|\mathbf{N}\rho_1|}$, where $C_1 = \sqrt{\varepsilon_1}$ (see §2). Thus size $\rho_i = O(S + \log \varepsilon_1)$, without fixing i. Moreover, $\varepsilon_1 < d^{\sqrt{|d|}}$ (cf. Hua [10]), thus $\log(\varepsilon_1) = O(\sqrt{|d|}\log|d|)$.* We will also need the constant C_1 in the following paragraphs.

It is now easy to analyze the complexity of step 3. We assume that the multiplications $\beta \leftarrow \beta \cdot \pi^{\min(e,f)}$ and divisions $\rho_{i-1} \leftarrow \rho_{i-1}/\pi^{\min(e,f)}$ are replaced by sequences of multiplications $\beta \leftarrow \beta \cdot \pi$ and divisions $\rho_{i-1} \leftarrow \rho_{i-1}/\pi$, which does not reduce the complexity. Then, apart from the final multiplication $\delta \leftarrow \beta\rho_{i-1}$, step 3 consists of a sequence of divisions and multiplications,

* Unfortunately, by the Brauer-Siegel theorem [16], Chapter XVI, not a much better estimate for $\log(\varepsilon_1)$ is to be expected since $h(d)\log(\varepsilon_1) \gg d^{1/2-\varepsilon}$ for $d \to \infty$.

where one of the operands has size $O(S + \sqrt{|d|}\log d)$ while the other is a prime factor of a rational integer $l < \sqrt{|d|}$. Such a prime factor is either a rational prime $p < \sqrt{|d|}$, or an element $\pi \in O_d$ with $|\mathbf{N}\pi| < \sqrt{|d|}$. In the latter case π will also have been adjusted by multiplication by an appropriate power of ε_1, so that size(π) = $O(\log|d| + \log\varepsilon_1) = O(\sqrt{|d|}\log d)$. Thus each operation requires $O((S + \sqrt{|d|}\log d)\sqrt{|d|}\log d)$ steps. Notice that whenever a division α/ψ is performed, $\mathbf{N}\psi$ must be computed, which takes only $O(d(\log d)^2)$ binary steps, since ψ will always be the second type of operator, with size $O(\sqrt{|d|}\log d)$. The number of multiplications and divisions to be performed in step 3 is easily seen to be $O(z)$, where z is the sum of the number of prime factors of ξ, η, and ρ_{i-1}, which is $O(S)$. In the final multiplication $\delta \leftarrow \beta\rho_{i-1}$ both factors have once again size $O(S + \sqrt{|d|}\log d)$. Thus the total complexity of all operations in step 3 becomes

$$O(S(S + \sqrt{|d|}\log d)\sqrt{|d|}\log d + (S + \sqrt{|d|}\log d)^2) = O(S(S + \sqrt{|d|}\log d)\sqrt{|d|}\log d).$$

It remains to analyze the complexity of step 2.

If we compute the exact representation of x_i and y_i, then this calculation would take $O(S^2)$ steps for a fixed d, which is too much to guarantee the desired overall complexity for all iterations. However, it suffices to compute floating point approximations for x_i, y_i, with an appropriate upper bound for the *absolute* error, 0.01 say. Then instead of $|\mathbf{N}\rho_{i+1}| < |\mathbf{N}\rho_i|/2$ we will accomplish $|\mathbf{N}\rho_{i+1}| < C_2|\mathbf{N}\rho_i|$ for some constant C_2, $1/2 < C_2 < 1$. This does not affect the analysis as far as asymptotic step count and bit sizes are concerned. Notice that the floating point approximations thus are local to step 2, and no error gets propagated to the next iteration. The exact values of x_i, y_i are

$$x_i = 2\,\frac{R\rho_{i-1}R\rho_i \pm I\rho_{i-1}I\rho_i}{\mathbf{N}\rho_i},$$

with the +sign applying for $d < 0$,

$$y_i = 2\,\frac{-R\rho_{i-1}I\rho_i + I\rho_{i-1}R\rho_i}{\mathbf{N}\rho_i\sqrt{|d|}}.$$

Let $A_i = \log\sqrt{|\mathbf{N}\rho_{i-1}|} - \log\sqrt{|\mathbf{N}\rho_i|}$. The numerators of x_i and y_i have absolute values $\le 2C_1^2\sqrt{|\mathbf{N}\rho_{i-1}\|\mathbf{N}\rho_i|}$. Hence there is a constant C_3 such that $\log|x_i|$, $\log|y_i| \le A_i + C_3$, and a constant C_4 such that the desired accuracy for x_i, y_i is guaranteed, if $A_i + C_4$ significant digits of both numerator and denominator of x_i and y_i are computed, and floating point division is applied.

For the following sufficiently exact error analysis we need to introduce a variant of the definition of size(α) $(\alpha \in O_d)$ given in the introduction. (The previous definition is insufficient because size$(a + a\sqrt{d})$ can be roughly 2 size(a) $(a \in \mathbf{Z})$, whereas we will need a bound for size$(a + a\sqrt{d}) -$ size(a) that is independent of a.) Define

$$\text{size}_1(\alpha) = \max(\log|R\alpha|,\ \log|I\alpha|).$$

Now the relation $|R(\rho_j)|, |I(\rho_j)| \le C_1\sqrt{|\mathbf{N}\rho_j|}$ also allows to find another constant C_5 such that

$$\left|\text{size}_1(\rho_j) - \log\sqrt{|\mathbf{N}\rho_j|}\right| \le C_5. \tag{4.2}$$

Therefore at most $2C_5$ leading digits can cancel out when the subtraction in the calculation of $\mathbf{N}\rho_j$ is performed. Applying (4.2) to $j = i-1$ and to $j = i$, A_i can be estimated by $(\text{size}_1\ \rho_{i-1} - \text{size}_1\ \rho_i)$ within an error bounded by a constant bound C_6; so the number of digits that ultimately have to be used is $(\text{size}_1\ \rho_{i-1} - \text{size}_1\ \rho_i) + C_4 + 2C_5 + C_6 = B$. Note that B, contrary to A_i, is known before x_i, y_i are computed. As pointed out, $C_1 = O(\sqrt{\varepsilon_1})$. Then it follows that C_3 ,…, C_6 are all of size $O(\log \varepsilon_1)$, which is $O(\sqrt{|d|}\log d)$, as stated above. Hence $B = O(A_i + \sqrt{|d|}\log d)$. All arithmetic with B-digit-precision can be performed in $O(B^2) = O((A_i + \sqrt{|d|}\log d)^2) = O(A_i^2 + |d|(\log d)^2)$ steps. The total complexity of all these operations for all iterations of step 2 to is therefore

$$O(\sum_{i\ge 1} A_i^2 + S|d|(\log d)^2) = O(S(S + |d|(\log d)^2)).$$

Floating point arithmetic only applies to x_i and y_i, not to other intermediate results like p_j, q_j $(j = 1,\ldots,n)$, x_{i+1} etc. We now consider the complexity of the remaining operations in an iteration of step 2. The continued fraction approximation p_n/q_n of y_i can be calculated as follows: from the B-digit floating point representation of y_i we obtain a fractional representation s/t, s, $t \in \mathbf{Z}$, where s has B digits and t is a power of 2. Note that $|y_i|$ cannot be larger than its B-digit mantissa s, since B was large enough to guarantee an absolute error < 1. If $|y_i| < 1/\sqrt{|d|}$, then instead of computing continued fractions we may simply choose $l_{i+1} = 1$, $y_{i+1} = 0$, which satisfies the required inequality $|l_{i+1}y_i - y_{i+1}| < 1/\sqrt{|d|}$. If not, then both $\log(s)$ and $\log(t)$ are $O(A_i + \sqrt{|d|}\log d)$. If g_1, g_2 are the quotients of the Euclidean algorithm for rational integers with non-negative remainders, applied to s, t, then the numbers p_i, q_i are determined by the following recurrence relation:

$$p_{-1} = q_0 = 0, \quad q_{-1} = p_0 = 1,$$

$$\left.\begin{aligned} p_i &= p_{i-2} + g_i p_{i-1} \\ q_i &= q_{i-2} + g_i q_{i-1} \end{aligned}\right\} \quad i = 1, 2, \cdots.$$

From these equations it becomes clear that $\sum_{i=1}^{n} \log(g_i) = O(\log q_n) = O(\log\sqrt{|d|})$, and since the intermediate results of the Euclidean algorithm applied to s, t, as well as the numbers p_i, q_i have $O(A_i + \sqrt{|d|}\log d)$ digits, we find by the usual method that $O((\log d)(A_i + \sqrt{|d|}\log d))$ binary steps are needed for the calculation of p_n, q_n.

The logarithms of rational and irrational part of γ_{i+1} are $O(A_i + \sqrt{|d|}\log d)$, since γ_{i+1} is an approximation of $l_{i+1}\rho_{i-1}/\rho_i$. This ensures the desired complexity bound $O((A_i + \sqrt{|d|}\log d)(S + \sqrt{|d|}\log d))$ for the multiplication $\gamma_{i+1}\rho_i$ in the computation of ρ_{i+1}. Then the same bound for the order of magnitude applies to the logarithms of x_{i+1}, y_{i+1}, p_1 ,…, p_n. The only step which

requires further consideration is the division of ρ_{i+1} by an appropriate power of ε_1.

Consider first the value $\rho_{i+1}^{(0)}$ of ρ_{i+1} before this adjustment. Recall that $\rho_{i+1}^{(0)} = \rho_i \tau$, where $|R\tau|$, $|I\tau| < 1$ (actually $\leq 1/2$ if the exact values of x_i, y_i are used). This implies $\text{size}_1 \rho_{i+1}^{(0)} \leq \text{size}_1 \rho_i + C_7$ for some constant C_7 with $C_7 = O(\log|d|)$. Let $\rho_{i+1}^{(t)}$ be the value of ρ_{i+1} after the adjustment. Then $\phi(\rho_{i+1}^{(0)}) = \phi(\rho_{i+1}^{(t)}) + k\phi(\varepsilon_1) \geq (k-1)\phi(\varepsilon_1)$. Together with the defining equality for the argument function ϕ, this implies $k = O(A_i)$. We can proceed as follows: first, if $R\rho_{i+1}^{(0)} < 0$, replace $\rho_{i+1}^{(0)}$ by $-\rho_{i+1}^{(0)}$. Then if $I\rho_{i+1}^{(0)} > 0$, compute $\rho_{i+1}^{(j)} = \rho_{i+1}^{(j-1)}/\varepsilon_1$, $j = 1, 2,...$ until $I\rho_{i+1}^{(j_0)} \leq 0$, so that $\phi\rho_{i+1}^{(j_0-1)} > 0$, $\phi\rho_{i+1}^{(j_0)} \leq 0$, then choose among $\rho_{i+1}^{(j_0-1)}$ and $\rho_{i+1}^{(j_0)}$ the one for which the rational part, or, equivalently, the absolute value of the irrational part, is smaller. If instead after the possible replacement of $\rho_{i+1}^{(0)}$ by $-\rho_{i+1}^{(0)}$ we get $I\rho_{i+1}^{(0)} < 0$, compute similarly $\rho_{i+1}^{(j)} = \rho_{i+1}^{(j-1)}\varepsilon_1$, $j = 1, 2, ...,$ until $I\rho_{i+1}^{(j_0)} \geq 0$. The number of divisions is $O(A_i)$, and each division has complexity $O((S + \sqrt{|d|}\log d)\sqrt{|d|}\log d)$.

We conclude that the highest asymptotic complexity within step 2 occurs for the adjustment of ρ_{i+1}, namely $O(A_i (S + \sqrt{|d|}\log d)\sqrt{|d|}\log d)$. Adding this up for $i = 1, 2, \ldots$, we obtain $O(S(S + \sqrt{|d|}\log d)\sqrt{|d|}\log d)$. We have established that this is indeed an upper bound for the complexity of all operations, including step 3. This completes the proof. □

Omitting the adjustment of ρ_{i+1} does not affect the correctness of the algorithm. In this case it is not obvious, however, that the complexity $O(S^2)$ can still be achieved for fixed d. Moreover, it is in any case desirable to carry out this normalization at least for the final result. After all, if $\text{GCD}(\xi, \eta) = 1$, it would be unsatisfactory to obtain some power $\pm\varepsilon_1^k$, $k \neq 0$, as output.

We now discuss how the preconditioning for this algorithm can be done. For the calculation of the fundamental unit ε_1 in the case $d > 0$, there exists the well-known algorithm using continued fractions. An improvement has been provided in Pohst & Zassenhaus [20]. The factorization of the multipliers $l < \sqrt{|d|}$ will begin by finding their factorization into rational primes p. Using the sieve of Eratosthenes, the largest prime factor for every l can be found in time $O(\sqrt{|d|})$, which essentially solves the problem. Then it can be determined efficiently, using the law of quadratic reciprocity, which of these rational primes p split further in O_d; see §5. One of the referees suggests the following algorithm for finding the prime factors π_1 and π_2, if p splits further. The idea is to explicitly find the transformation from the quadratic form $(1, 0, -d/4)$ or $(1, 1, (1-d)/4)$, whichever is integral, to a form $(4p, g_1, g_2)$, all of discriminant d. Consider first the case $d \equiv 0 \pmod 4$. Then the transforming matrix $M = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix}$ yields the solution $\alpha^2 - d\beta^2 = 4p$, the wanted factorization. For $d \equiv 1 \pmod 4$, the matrix M yields $(2\alpha + \beta)^2 - d\beta^2 = 4p$ as the solution. In order to compute M one appeals to the theory of reduction on the principal cycle of reduced quadratic forms of discriminant d (cf [26],. Section 4, pp. 256-261). One starts at a form $(4p, g_1, g_2)$ or (p, g_1, g_2) of discriminant d, which is easily

determined using an algorithm for taking square roots modulo p. Then one steps through the principal cycle by repeated reduction until one detects the forms $(1, 0, -d/4)$ or $(1, 1, (1-d)/4)$, respectively. For each prime p the running time depends on the number of forms in in the cycle, about $O(d^{1/2+\varepsilon})$ such that the procedure for all primes $< \sqrt{|d|}$ takes $O(d^{1+\varepsilon})$.

We conclude by briefly discussing the application of this algorithm to domains O_d with $h > 1$. The algorithm itself does not change, except that prime ideal factors of (l) for the multipliers l have to be considered, rather than prime factors of l in O_d. What does it accomplish? Of course it would be a trivial task to find a representation for the greatest common divisor of the ideal (ξ) and (η), as (ξ, η) would already be such a representation. Instead, it is desirable to put it into some normal form. Let $\{I_1, \ldots, I_h\}$ be a set of representatives of the ideal classes. As is well-known, the I_i can be chosen such that

$$||I_i|| \leq \begin{cases} \dfrac{\sqrt{d}}{2} & \text{for } d > 0, \\ \dfrac{2\sqrt{|d|}}{\pi} & \text{for } d < 0, \end{cases}$$

where $||I||$ is the number of congruence classes mod I. Then every ideal of O_d has a unique representation of the form $(\alpha)/I_i$, $\alpha \in O_d$, $1 \leq i \leq h$. The GCD algorithm can now be used to compute the normal form of (ξ, η). The following preparatory work is needed: one has to compute the normal forms of the ideal prime factors of l, $l = 2, \ldots, \lfloor \sqrt{|d|} \rfloor$ and a multiplication table for the normal forms of the ideal products $I_i \cdot I_j$. The latter problem is equivalent to the computation of the class group and can be done in time $|d|^{1/4+o(1)}$, see Schoof [26]. The factorization of the ideals (l) is at least as difficult as in the case $h(d) = 1$. After this preparatory work has been finished, the complexity bound of theorem 4.1 applies again.

5. A Constructive Version of a Theorem by A. Thue with Application

A theorem by Axel Thue [30] states that if a and m are positive relatively prime integers then there exist integers $0 < x \le \sqrt{m}$, $0 < |y| \le \sqrt{m}$ such that $ax + y \equiv 0 \bmod m$. This theorem and its generalizations (cf. Brauer and Reynolds [2] and Nagell [19]) are usually proved using the pigeon hole principle. The following theorem shows how *all* solutions for the above congruence can be found in $\log^2 m$ steps.

Theorem 5.1: Let $a \ge 1$, m, e, $f \ge 2$ be integers such that $a < m$, $(e-1)(f-1) < m < ef$. Then the problem

$$m \mid ax + y, \quad 0 < x < e, \quad |y| < f, \quad y \ne 0 \tag{5.1}$$

is solvable in integers x, y if and only if $d = \mathrm{GCD}(a, m) < f$. Furthermore, assuming that this is the case, let

$$\frac{p_0}{q_0} = \frac{0}{1}, \frac{p_1}{q_1}, \ldots, \frac{p_N}{q_N} = \frac{a/d}{m/d}, \quad q_N \ge \frac{m}{f-1} > e - 1,$$

be the continued fraction approximations of a/m and choose n such that $q_n < e \le q_{n+1}$. Then $x_1 = q_n$, $y_1 = mp_n - aq_n$ is a solution for (5.1). The set of all solutions for (5.1) exclusively either consists of λx_1, λy_1, $1 \le \lambda < \min(e/x_1, f/|y_1|)$ or else consists of x_1, y_1 and x_2, y_2 with $y_1 y_2 < 0$. In the latter case we can determine x_2, y_2 from p_{n-1}/q_{n-1} or p_{n+1}/q_{n+1} in $O(\log^2 m)$ steps.

Before we can prove this theorem we need to establish a lemma from the theory of continued fraction approximations. Following Hardy and Wright [8], §10, we denote a continued fraction by

$$[a_0, a_1, \ldots, a_n, \cdots] = a_0 + \frac{1}{[a_1, \ldots, a_n, \cdots]}.$$

The nth convergent is given by

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$$

and satisfies $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$. Notice that

$$[a_0, a_1, \ldots, a_n, 1] = [a_0, a_1, \ldots, a_n + 1] \tag{5.2}$$

but this is the only ambiguity possible for the *simple* continued fraction expansion of a real number where a_i, $i > 0$, are positive integers.

Lemma 5.1 ([8], Theorem 172): If

$$x = \frac{PZ + R}{QZ + S},$$

where $Z \ge 1$ and P, Q, R, and S are integers such that

$$Q > S > 0, \quad PS - QR = \pm 1,$$

then R/S and P/Q are two consecutive convergents in the simple continued fraction expansion of x provided we choose the LHS of (5.2) to resolve ambiguity.

Proof: Consider the continued fraction expansion

$$\frac{P}{Q} = [a_0, a_1, \ldots, a_n] = \frac{p_n}{q_n}.$$

From (5.2) it follows that we may choose n even or odd as we need it. Now let n be such that

$$PS - QR = (-1)^{n-1} = p_n q_{n-1} - p_{n-1} q_n.$$

Now GCD(P, Q) = 1 and $Q > 0$ and hence $P = p_n$ and $Q = q_n$ and therefore

$$p_n (S - q_{n-1}) = q_n (R - p_{n-1}).$$

This implies that $q_n \mid (S - q_{n-1})$ which by virtue of $q_n = Q > S > 0$ and $q_n > q_{n-1} > 0$ is only possible if $S - q_{n-1} = 0$. Therefore $R = p_{n-1}$ and $S = q_{n-1}$ and

$$x = \frac{p_n Z + p_{n-1}}{q_n Z + q_{n-1}} \text{ implies } x = [a_0, a_1, \ldots, a_n, Z]. \quad \square$$

Proof of Theorem 5.1: Since $m \mid ax + y$ there exists an integer z such that $y = mz - ax$ and thus $d \mid y$ which implies $d \le |y| < f$. For p_n/q_n we have, as in the proof of lemma 4.1,

$$\left| \frac{a}{m} - \frac{p_n}{q_n} \right| \le \frac{1}{q_n q_{n+1}} \le \frac{1}{q_n e},$$

therefore $|y_1| = |aq_n - mp_n| \le m/e < f$. Notice that $y_1 \ne 0$ because $a/m \ne p_n/q_n$ for $q_n < q_N = m/d$.

We prove next that if $\hat{x}_1$, $\hat{y}_1$ also solve (5.1) and $y_1 \hat{y}_1 > 0$ then $y_1/x_1 = \hat{y}_1/\hat{x}_1$. Since $\hat{x}_1 (ax_1 + y_1) \equiv x_1 (a\hat{x}_1 + \hat{y}_1) \equiv 0 \bmod m$, $y_1 \hat{x}_1 - \hat{y}_1 x_1 \equiv 0 \bmod m$. But $|y_1 \hat{x}_1|, |\hat{y}_1 x_1| \le (e-1)(f-1) < m$, hence $y_1 \hat{x}_1 = \hat{y}_1 x_1$, which is our claim. Nonetheless it can happen that GCD(x_1, y_1) = $g \ne 1$ and therefore the assertion $\hat{x}_1 = \lambda x_1$, $\hat{y}_1 = \lambda y_1$ needs proof. First we note that

$$ax_1 + y_1 = mp_n \text{ and } \mathrm{GCD}(p_n, x_1) = \mathrm{GCD}(p_n, q_n) = 1.$$

Thus

$$a\lambda \frac{x_1}{g} + \lambda \frac{y_1}{g} = \frac{\lambda m p_n}{g} \equiv 0 \bmod m$$

if and only if $g \mid \lambda$, since GCD(g, p_n) = 1.

Assume now that there exists a second solution x_2, y_2 to (5.1) such that $y_1 y_2 < 0$. Let $ax_1 + y_1 = mz_1$, $ax_2 + y_2 = mz_2$. Again by multiplying the first equation with x_2, the second with x_1

and subtracting we get $x_1y_2 - x_2y_1 \equiv 0 \bmod m$. Since $|x_1y_2|, |x_2y_1| < m$ and $y_1y_2 < 0$ we must have $|x_1y_2 - x_2y_1| = m$. Thus $m\,|x_1z_2 - x_2z_1| = |x_1y_2 - x_2y_1| = m$ which implies that

$$|x_1z_2 - x_2z_1| = 1. \tag{5.3}$$

One immediate conclusion from (5.3) is that no solutions proportional to either x_1, y_1 or x_2, y_2, the only other possible solutions (as shown before), can occur. For otherwise, for any of these solutions, say $\hat{x}$, $\hat{y}$, GCD($\hat{x}$, $\hat{z}$) $\neq 1$ where $\hat{z}$ is the corresponding, also proportional, multiplier of m.

It is harder to show how this second solution x_2, y_2 can be computed in case it exists. Surprisingly, this alternate solution can arise in two different ways. Without loss of generality let us assume that $x_2 < x_1$. Notice that by this assumption we now only know that either x_1, y_1 or x_2, y_2 is the solution found as stated in the theorem.

Case 1: $|y_2| > |y_1|$. Let $Z = |y_2/y_1| > 1$. Then $Zmz_1 = Zax_1 + Zy_1 = Zax_1 - y_2$, hence $Zmz_1 + mz_2 = Zax_1 + ax_2$, or

$$\frac{a}{m} = \frac{Z\,z_1 + z_2}{Z\,x_1 + x_2}.$$

All conditions to lemma 5.1 with $P = z_1$, $Q = x_1$, $R = z_2$, and $S = x_2$ are now satisfied (refer in particular to (5.3)) and we can conclude that z_2/x_2 and z_1/x_1 must be consecutive convergents to a/m. Therefore $x_2 = q_{n-1}$ and $y_2 = mp_{n-1} - aq_{n-1}$.

Case 2: $|y_2| \le |y_1|$. Consider for an integer $k \ge 0$,

$$x_3 = x_1 + kx_2, \quad y_3 = y_1 + ky_2, \quad z_3 = z_1 + kz_2,$$

such that

$$|y_2| \ge |y_3| \quad \text{and} \quad \operatorname{sign}(y_3) = -\operatorname{sign}(y_2).$$

Now

$$0 < x_2 < x_3, \;\; \frac{a}{m} = \frac{Zz_3 + z_2}{Zx_3 + x_2}, \;\; Z = \left|\frac{y_2}{y_3}\right| \ge 1, \;\; z_3x_2 - z_2x_3 = \pm 1,$$

and lemma 5.1 applies again. Therefore z_2/x_2 and z_3/x_3 are consecutive continued fraction approximations. Notice that $Z = 1$ is possible if and only if z_2/x_2 is the second to last convergent of a/m. In that case

$$\frac{z_3}{x_3} = [0, a_1, \ldots, a_N - 1] \quad \text{where} \quad \frac{a}{m} = [0, a_1, \ldots, a_N].$$

In order to compute x_1 and y_1 we must find k. Since $|y_1| = |y_3 - ky_2|$ is monotonicly increasing with k we choose the smallest k such that $x_3 - kx_2 \le e - 1$. In other words, the only possible value is

$$k = \left\lceil \frac{x_3 - e + 1}{x_2} \right\rceil .$$

Observe that $x_1 = x_3 - kx_2 \neq x_2$ since GCD(x_2, x_3) = 1 and

$$ax_1 + y_1 \equiv 0 \bmod m \quad \text{for} \quad y_1 = y_3 - ky_2. \quad \square$$

The following example shows that all three types of solutions, namely either a single one, or two, or a family of proportional solutions do occur.

Example: Let $m = 11$, $e = f = 4$, $a = 7$. The continued fraction expansion of 7/11 is [0, 1, 1, 1, 3] and the convergents are 0/1, 1/1, 1/2, 2/3, and 7/11. Hence $x_1 = 3$, $y_1 = 2\cdot 11 - 7\cdot 3 = 1$ and $x_2 = 2$, $y_2 = 1\cdot 11 - 7\cdot 2 = -3$ are the only solutions for (5.1) in this case. For $a = 2$ the only solution is $1\cdot 2 - 2 \equiv 0 \bmod 11$ but for $a = 1$ there are three solutions, $1\cdot 1 - 1 \equiv 2\cdot 1 - 2 \equiv 3\cdot 1 - 3 \equiv 0 \bmod 11$.

Next let $m = 244$, $a = 47$, $e = 7$, and $f = 39$. The first three convergents are 0/1, 1/5, and 5/26. Thus $x_2 = 5$ and $y_2 = 244\cdot 1 - 47\cdot 5 = 9$. The second solution is obtained from $k = \lceil (26 - 7 + 1)/5 \rceil = 4$ as $x_1 = 26 - 4\cdot 5 = 6$ and $y_1 = -38$.

Finally let $m = 56$, $a = 21$, $e = 7$, and $f = 9$. The continued fraction expansion of 21/56 = [0, 2, 1, 2] and the convergents are 0/1, 1/2, 1/3, and 3/8. Thus $x_2 = 3$, $y_2 = -7$. We obtain z_3/x_3 = [0, 2, 1, 1] = 2/5, $k = \lceil (5 - 7 + 1)/3 \rceil = 0$, $x_1 = 5$, and $y_1 = 7$. $\square$

One application of theorem 5.1, described by P. Wang [31] and others, is to recover rational numbers from their modular representations. Set $y/x \in \mathbf{Q}$ and suppose we have found bounds e and f for the denominator and numerator, respectively. Then having computed $a = y\,x^{-1} \bmod m$, GCD(x, m) = 1, $(e-1)(f-1) < m$, we can find y/x by continued fraction approximation. Unfortunately, we may get two possible fractions y_1/x_1, y_2/x_2 of opposite sign. One way to resolve the ambiguity is to choose the e twice the bound of the denominator and select the solution with $x < e/2$, $|y| < f$. Wang, in fact, chooses m such that $\lceil \sqrt{m/2} \rceil$ is a bound for both numerator and denominator. In this application the existence of a solution is assumed and Thue's theorem does not come into play.

We now apply theorem 5.1 to the problem of factoring a rational prime p in the UFD O_d, d fixed. We again must precondition our algorithm by factoring all rational primes smaller than $2\sqrt{|D|}$ ($\sqrt{D}$ suffices for $D > 0$). First consider $p \nmid d$. From §2 we know that $p > \sqrt{D}$ factors if and only if $(D/p) = +1$. In that case a prime factor of p is $\pi = \mathrm{GCD}(p, l + \sqrt{D})$ where $l^2 \equiv D$ modulo p. We can compute l by either the Tonelli-Shanks algorithm (cf D. Knuth [14], Sec. 4.6.2, Exercise 15) or by R. Schoof's [27] algorithm. The latter is deterministic and runs in $O(\log^8 p)$ steps, since $|d|$ is fixed. The GCD algorithm of §3 can give us the wanted factor π but in this special case theorem 5.1 can be applied to our advantage. Let $e = \left\lceil \sqrt{p/\sqrt{|D|}} \right\rceil$ and $f = \left\lceil \sqrt{p\sqrt{|D|}} \right\rceil$. Then $(e-1)(f-1) < p < ef$ and $2 \le e \le f$ for $p > \sqrt{|D|}$. Then we compute the continued fraction

approximation to l/p and get integers x, y such that $p \mid yl + x$, $0 < y < e$, $|x| < f$. This solution satisfies

$$0 \equiv (x + l\,y)(x - l\,y) \equiv x^2 - D\,y^2 \bmod p,$$

thus there exists an integer q with $x^2 - D\,y^2 = q\,p$. By our bounds we get

$$x^2 - D\,y^2 < p\sqrt{|D|} - D\,\frac{p}{\sqrt{|D|}} = 2\sqrt{|D|}\;p \text{ for } D < 0,$$

$$|x^2 - D\,y^2| < \max(x^2, D\,y^2) < \sqrt{D}\;p \text{ for } D > 0,$$

thus $|q| < 2\sqrt{|D|}$.

In the other case $p \mid d$, $p > \sqrt{|D|}$, we have $x^2 - Dy^2 = qp$ with $x = 0$, $y = 1$ and $q < \sqrt{D}$. In both cases the factorization of q into primes $q = \gamma_1 \cdot \cdot \cdot \gamma_k$ in O_d is already known. Since O_d is a UFD and

$$(x + \sqrt{D}\,y)\,(x - \sqrt{D}\,y) = \gamma_1 \cdot \cdot \cdot \gamma_k\;p,$$

γ_i must divide $x + \sqrt{D}\,y$ or $x - \sqrt{D}\,y$. Let γ be a maximum product of γ_i such that γ divides $x + \sqrt{D}\,y$. Then q/γ divides $x - \sqrt{D}\,y$ and the prime factors of p are then $(x + \sqrt{D}\,y)/\gamma$ and $(x - \sqrt{D}\,y)/(q/\gamma)$. To prove this, we only have to show that neither quotient is a unit. This follows from the fact that the division can only decrease the norm of the dividend. If one quotient became a unit the other one would have to have the norm of their product, p^2, which is larger than its original norm qp.

We now discuss how to factor $\xi \in O_d$ with O_d a unique factorization domain. We first factor $\mathbf{N}\,\xi = p_1 \cdot \cdot \cdot p_k$ over the integers. If $(d/p_i) = +1$ or p_i divides d we split $p_i = \pi_i\,\bar{\pi}_i$ by the algorithm discussed above. We thus obtain a factorization $\xi\,\bar{\xi} = \pi_1 \cdot \cdot \cdot \pi_l$ and it remains to trial divide ξ by π_i, $1 \le i \le l$, to determine which are its prime factors.

6. Conclusion

We have described algorithms for taking the greatest common divisor and computing the prime factorization of numbers in quadratic fields with unique factorization. The methods also apply to computing canonical representations of unions of ideals in quadratic number rings without unique factorization. Our algorithms are of polynomial running time provided we fix the discriminant. We have also shown how to reduce factorization in quadratic number rings with unique factorization to rational integer factorization. If the discriminant is large, say of order 10^{10}, our algorithms unfortunately become impractical. Future investigations will focus on how to treat these cases efficiently.

Acknowledgement: We thank Andrew Odlyzko for providing us with many references to the literature. The referee has scrupulously read and corrected two versions of the manuscript, as well as suggested many improvements and several references unknown to us. For that we are very grateful.

Note added September 22, 2006: corrected last line of proof of Theorem 5.1: replaced $y_1 = y_3 - kx_2$ by $y_1 = y_3 - ky_2$.

References

1. Barnes, E. S. and Swinnerton-Dyer, H. P. F., "The homogeneous minima of binary quadratic forms," *Acta Math.*, 87, pp. 255-323 (1952).

2. Brauer, A. and Reynolds, R. L., "On a theorem of Aubry-Thue," *Canadian J. Math.*, 3, pp. 367-374 (1951).

3. Caviness, B. F. and Collins, G. E., "Algorithms for Gaussian integer arithmetic," *Proc. 1976 ACM Symp. Symbolic Algebraic Comp.*, pp. 36-45 (1976).

4. Chatland, H. and Davenport, H., "Euclid's algorithm in real quadratic fields," *Canadian J. Math.*, 2, pp. 289-296 (1950).

5. Cohn, P. M., "On the structure of $\mathbf{GL}_2$ of a ring," *Publ. Math. IHES*, 30, pp. 365-413 (1966).

6. Cooke, G. E., "A weakening of the Euclidean property for integral domains and applications to algebraic number theory I," *J. reine angew. Math.*, 282, pp. 133-156 (1976).

7. Cooke, G. E. and Weinberger, P. J., "On the construction of division chains in algebraic number rings, with applications to SL_2," *Comm. Algebra*, 3, 6, pp. 481-524 (1975).

8. Hardy, G. H. and Wright, E. M., *An Introduction to the Theory of Numbers*, Oxford Univ. Press, Oxford (1979).

9. Hasse, H., *Vorlesungen über die Zahlentheorie*, Springer Verlag, Berlin (1950). In German.

10. Hua, L.-K., "On the least solution of Pell's equation," *Bull. Amer. Math. Soc.*, 48, pp. 731-735 (1942).

11. Hurwitz, A., *Mathematische Werke*, 2, Birkhäuser Verlag, Basel (1963).

12. Kannan, R., "Algebraic geometry of numbers" in *Annual Review in Computer Science*, ed. J. F. Traub, 2, pp. 231-67, Annual Reviews Inc., Palo Alto, California (1987).

13. Kannan, R. and Bachem, A., "Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix," *SIAM J. Comp.*, 8, pp. 499-507 (1981).

14. Knuth, D. E., *The Art of Programming, vol. 2, Semi-Numerical Algorithms, ed. 2*, Addison Wesley, Reading, MA (1981).

15. Lamé, G., "Note sur la limite du nombre des divisions dans la recherche du plus grand diviseur entre deux nombres entiers," *Comptes Rendus Acad. Sci. Paris*, 19, pp. 867-870 (1844).

16. Lang, S., *Algebraic Number Theorie*, Addison-Wesley Publ., Reading, Massachusetts (1970).

17. Lazard, D., “Le meilleur algorithme d’Euclide pour $K[X]$ et $\mathbf{Z}$,” *Comptes Rendus Acad. Sci. Paris*, 284, pp. 1-4 (1977).

18. Lenstra, A. K., Jr., H. W. Lenstra, and Lovász, L., “Factoring polynomials with rational coefficients,” *Math. Ann.*, 261, pp. 515-534 (1982).

19. Nagell, T., “Sur un theoreme d’Axel Thue,” *Arkiv för Matematik*, 1, 33, pp. 489-496 (1951).

20. Pohst, M. and Zassenhaus, H., “An effective number theoretic method of computing the fundamental units of an algebraic number field,” *Math. Comp.*, 31, pp. 745-770 (1977).

21. Rolletschek, H., “The Euclidean algorithm for Gaussian integers,” *Proc. EUROCAL ’83, Springer Lec. Notes Comp. Sci.*, 162, pp. 12-23 (1983).

22. Rolletschek, H., “On the number of divisions of the Euclidean algorithm applied to Gaussian integers,” *J. Symbolic Comp.*, 2, pp. 269-291 (1986).

23. Rolletschek, H., “Shortest division chains in imaginary quadratic number fields,” Manuscript, Kent State Univ. (December 1987).

24. Samuel, P., “About Euclidean rings,” *J. Algebra*, 19, pp. 282-301 (1971).

25. Schönhage, A., “Schnelle Kettenbruchentwicklungen,” *Acta Inf.*, 1, pp. 139-144 (1971). (In German).

26. Schoof, R. J., “Quadratic fields and factorization” in *Computational Methods in Number Theory, Part II*, ed. H. W. Lenstra, Jr. and R. Tijdeman, Mathematical Centre Tracts, 154, pp. 89-139, Mathematisch Centrum, Amsterdam (1982).

27. Schoof, R. J., “Elliptic curves over finite fields and the computation of square roots mod p,” *Math. Comp.*, 44, pp. 483-494 (1985).

28. Shanks, D., *Solved and Unsolved Problems in Number Theory I, 2nd ed.*, Chelsea Publ. (1978).

29. Stark, H., “A complete determination of complex quadratic fields of class number one,” *Mich. Math. J.*, 17, pp. 1-27 (1967).

30. Thue, A., “Et par antydninger til en talteoretisk methode,” *Vid. Selsk. Forhandlinger Christiania*, 7 (1902). In Norwegian.

31. Wang, P. S., “A p-adic algorithm for univariate partial fractions,” *Proc. 1981 ACM Symp. Symbolic Algebraic Comp.*, pp. 212-217 (1981).

32. Weinberger, P. J., “On Euclidean rings of algebraic integers” in *Proc. Symp. Pure Math.*, 24, pp. 321-332, Amer. Math. Soc. (1973).