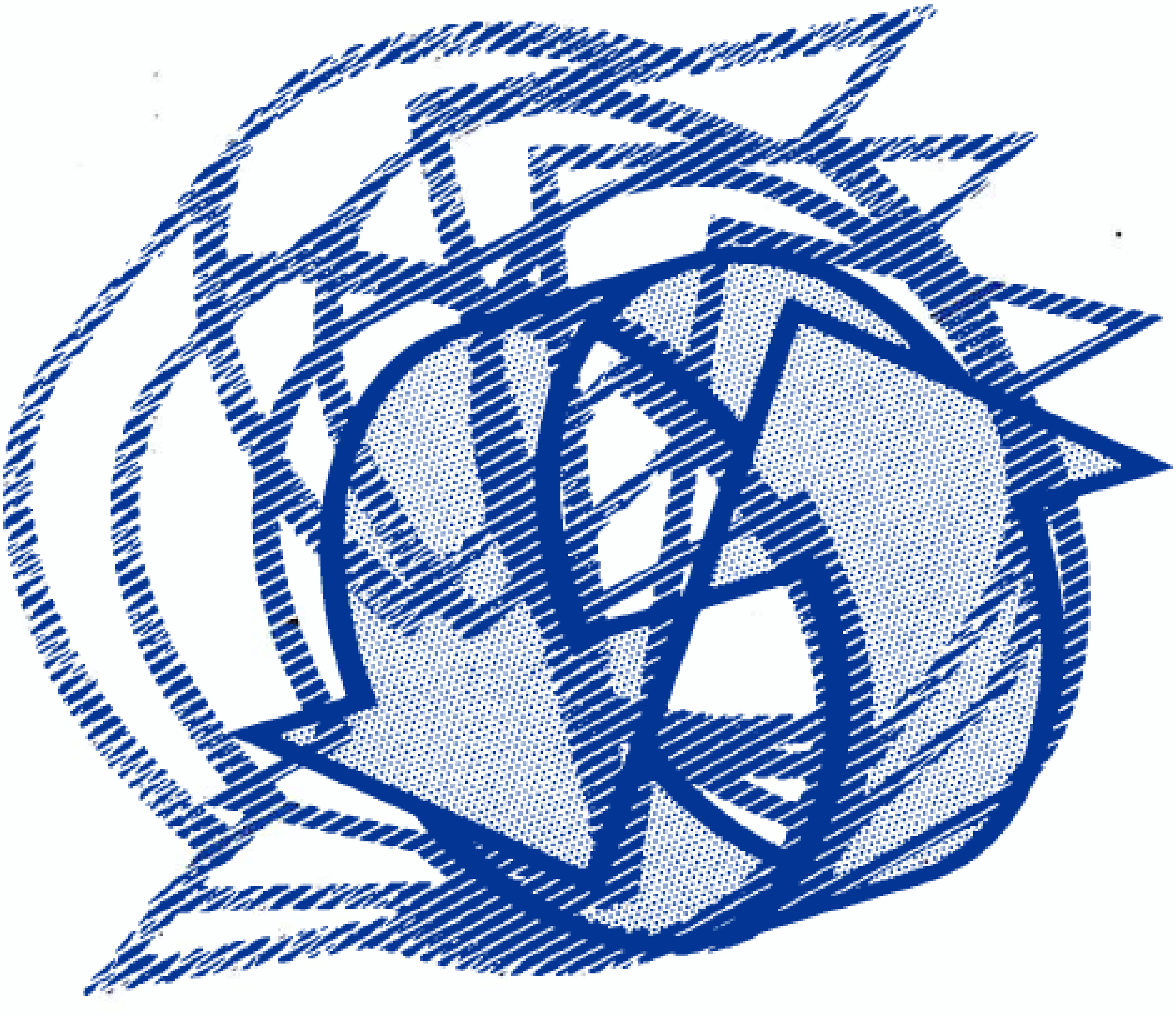
CWI Tract 65

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

Algorithms for diophantine equations

B.M.M. de Weger



CWI Tracts

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O).

CWI Tract 65

Algorithms for diophantine equations

B.M.M. de Weger

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

1980 Mathematics Subject Classification: 11Y50, 11D61.
ISBN 90 6196 375 3
NUGI-code: 811

Acknowledgements.

The research on which this book reports has been done while I worked for the Netherlands Foundation for Mathematics SMC, with financial support from the Netherlands Organization for the Advancement of Pure Research ZWO. This research took place from 1983 to 1987 at the University of Leiden, under supervision of Professor R. Tijdeman and Dr. F. Beukers.

I am very grateful to my (number theory) teacher, Professor R. Tijdeman (Leiden), for suggesting the research topic, for all his help, comments and criticism, on mathematics and everything else. I am also indebted to:
→ Dr. F. Beukers (Utrecht), for comments and discussions,
→ Dr. A. Pethö (Debrecen), my first coauthor, for the cooperation, the hospitality in Cologne in february 1985, and for allowing me to publish our joint work in Chapter 4 of this book,
→ Prof. N. Tzanakis (Iraklion), my other coauthor, for the cooperation, the many discussions, the hospitality in Iraklion in october-november 1986, for allowing me to publish our joint work in Chapter 8 of this book, and for pointing out some errors in the manuscript,
→ Prof. L. Wang (Beijing), for carefully checking most of the computations of Chapter 6, and thus finding some errors,
→ Dr. B.H. Gilding (Enschede), for polishing some of the english,
→ the Faculty of Applied Mathematics of the University of Twente (Enschede), for providing a good working environment and computing and text-editing facilities,
→ the Dutch Open University (Heerlen), for (unintentionally) providing text-editing facilities,
and finally to Alda, for being there and loving me.

SOLI DEO GLORIA.

Benne de Weger,
University of Twente,
Enschede, The Netherlands.

February 1989.

Contents.

Chapter 1. Introduction.

1.1. Algorithms for diophantine equations.

This monograph deals with certain types of *diophantine equations*. An *equation* is a mathematical formula, expressing equality of two expressions that involve one or more unknowns (variables). *Solving* an equation means finding all *solutions*, i.e. the values that can be substituted for the unknowns such that the equation becomes a true statement. An equation is called a *diophantine* equation if the solutions are restricted to be *integers* in some sense, usually the ordinary rational integers (elements of $\mathbb{Z}$) or some subset of that.

Examples of diophantine equations that will be studied in this book are

$$x^2 + 7 = 2^n$$

(the Ramanujan-Nagell equation, having only the solutions given by $(\pm x,n) = (1,3), (3,4), (5,5), (11,7), (181,15)$, see Chapter 4);

$$2^x = 3^y + 5^z$$

(a purely exponential equation, having only the solutions $(x,y,z) = (1,0,0), (2,1,0), (3,1,1), (5,3,1), (7,1,3)$, see Chapter 6);

$$y^2 = x^3 - 4\cdot x + 1$$

(an elliptic curve equation, having only 22 solutions, of which the largest are $(x,y) = (1274,\pm 45473)$, see Chapter 8). The three examples mentioned here are only some examples; we will study much wider classes of equations. We also study (in Chapter 5) a *diophantine inequality* (a formula expressing that one expression is larger than another, where solutions are again restricted to integers). In the following discussion the statements about diophantine equations also hold for this inequality.

What the equations treated in this book have in common is that they can all be solved by the same method. This method consists essentially of three parts: a transformation step, an application of the Gelfond-Baker theory, and a diophantine approximation step. We explain these steps briefly.

To start with, one transforms the equation into a purely exponential equation or inequality, i.e. a diophantine equation or inequality where the unknowns are all in the exponents, such as in the second example given above. Each type of diophantine equation needs a particular kind of transformation, so that it is difficult to be more specific at this point. In some instances, such as in the second example above, this transformation is easy, if not trivial. In other instances, as in the first example above, it uses some arguments from algebraic number theory, or, as in the third example above, a lot of them.

In general, such a purely exponential equation has the form

$$\sum_{i=1}^{t} c_i \cdot \prod_{j=1}^{s_i} \alpha_{ij}^{n_{ij}} = c_0 \cdot \prod_{j=1}^{s_0} \alpha_{0j}^{n_{0j}} , \tag{1.1}$$

and a corresponding purely exponential inequality looks like

$$\left| \sum_{i=1}^{t} c_i \cdot \prod_{j=1}^{s_i} \alpha_{ij}^{n_{ij}} \right| < \min_i \left| c_i \cdot \prod_{j=1}^{s_i} \alpha_{ij}^{n_{ij}} \right|^{\delta} \tag{1.2}$$

where t, s_i, c_i, α_{ij}, δ are constants with $t, s_i \in \mathbb{N}$, $0 < \delta < 1$, and c_i, α_{ij} belong to some algebraic extension of $\mathbb{Q}$, and where the n_{ij} are the unknowns in $\mathbb{Z}$. We now suppose that the number of terms t on the left hand side of (1.1) or (1.2) is equal to 2 . This restriction is essential for the second step, in which we use results from the so-called theory of linear forms in logarithms, also known as the Gelfond-Baker theory. (Some special exponential equations of type (1.1) with $t > 2$ can also be treated by the Gelfond-Baker method, since they can be reduced to exponential inequalities of type (1.2) with $t = 2$, cf. Stroeker and Tijdeman [1982], Alex [1985[a]], [1985[b]], Tijdeman and Wang [1988].)

An exponential equation or inequality such as (1.1) or (1.2) with $t = 2$ gives rise to a *linear form in logarithms*

$$\Lambda = \log \beta_0 + \sum_{i=1}^{m} n_i \cdot \log \beta_i ,$$

where the β_i are algebraic constants, and the n_i are integral unknowns. Here, the logarithms are real or complex in some instances, or p-adic in other cases. This relation between equation and linear form in logarithms is such that for a large solution of the equation the linear form is extremely close to zero (in the real or complex sense, or in the p-adic sense). The Gelfond-Baker theory provides effectively computable lower bounds for the absolute values (respectively p-adic values) of such linear forms in logarithms of algebraic numbers. In many cases these bounds have been explicitly computed. Comparing the so-found upper and lower bounds it is possible to obtain explicit upper bounds for the solutions of the exponential diophantine equation or inequality, leading to upper bounds for the solutions of the original equation. This second step, unlike the first (transformation) step, is of a rather general nature.

We remark that many authors have given effectively computable upper bounds for the solutions of a wide variety of diophantine equations, by applying the method sketched above. For a survey, see Shorey and Tijdeman [1986]. Often these authors were satisfied with the knowledge of the existence of such bounds, and they did not actually compute them. If they computed bounds, they did not always determine all the solutions. In this book, solving an equation will always mean: explicitly finding all the solutions.

After the second step, the problem of solving the diophantine equation is reduced to a finite problem, which is treated in the third part of the method. Namely, since we have found explicit upper bounds for the absolute values of the (integral) unknowns, we have to check only finitely many possibilities for the unknowns. However, the word *finite* does not mean the same as *small* or *trivial*. In fact, the constants appearing in the lower bounds that the Gelfond-Baker theory provides for linear forms in logarithms are rather large. Therefore, in practice the upper bounds that can be obtained in this way for the solutions of purely exponential equations can be for instance as large as 10^{40}. This is far too large to admit simple enumeration of all the possibilities, even with the fastest of computers today.

Proving the existence of an absolute upper bound for the solutions reduces the determination of all the solutions from an infinite task to a finite one. Thus, the application of the Gelfond-Baker theory (the second step) is in a sense infinitely many times as difficult a task than the only finite amount of checking that remains to be done (in the third step). Furthermore, this checking seems to be a technical problem only, not a mathematical one.

Nevertheless, it is the author's opinion that solving this comparatively small technical problem is not only nontrivial, but involves some serious and interesting mathematics. This book hopefully illustrates this opinion.

Notwithstanding the fact that the application of the Gelfond-Baker theory in the second step yields very large upper bounds, it is generally assumed that these upper bounds are far from the actual largest solution. Therefore, it is worthwile to search for methods to reduce these upper bounds to a size that can be more easily handled. Often it is possible to devise such a method using directly certain properties of the original diophantine equation, for example that large solutions must satisfy certain congruences modulo many or large numbers (Grinstead [1978], Brown [1985], Pinch [1988]), or some reciprocity condition (Pethö [1983]). The disadvantage of such methods is that they work only for that particular type of diophantine equation, so that in general for each type of equation a new reduction method must be devised. It would therefore be interesting to have methods for reducing upper bounds for the solutions of inequalities for linear forms in logarithms. They would be useful for solving any type of diophantine problem that leads to such inequalities.

Such methods are searched for in the third step of our method of solving diophantine equations. It is mainly in this third part that new developments can be reported. The arguments we use in the first and second parts are mainly classical, and we apply them to types of equations that have been studied before, and also to new types of equations.

The methods that are needed in the third step are provided by that part of the theory of *diophantine approximation* that is concerned with studying how close to zero a linear form can be for given values of the variables. Recently important progress has been made in this field, the breakthrough being the invention in 1981 by L. Lovász of the so-called L^3-laticce basis reduction algorithm. We will show how this L^3-algorithm leads to practically efficient *diophantine approximation algorithms*, which can be employed for many diophantine equations to show that in a certain interval $[X_1, X_0]$ no solutions exist. Usually X_1 is of the order of magnitude of $\log X_0$. When for X_0 the theoretical upper bound for the solutions is substituted, a new, and usually much better upper bound X_1 is found. For many equations the initial upper bound X_0 is well within reach of practical application of these algorithms, within only a few minutes of computer time. This thus leads in practice to methods for finding all the solutions of many types of diophantine equations, for which alternative methods have not yet been found or employed with success.

The method outlined above, and used in this book to solve many examples of various diophantine equations, is of an "algorithmic" nature. In a sense it lies between "ad hoc" methods and "theoretical" methods. This we shall explain below. Let a set of diophantine equations with an unspecified parameter in it be given. As an example of such a set, consider the generalized Ramanujan-Nagell equation $x^2 + D = 2^n$, where D is a parameter, and x, n are the unknowns.

An *ad hoc method* is a method for solving the equation for specific values of the parameters only. It may not work at all for other than these particular values. The first example of solving an equation of the type $x^2 + D = 2^n$ occurring in the literature is that by Nagell [1948] of $D = 7$. The method he used is of an ad hoc nature, since it depends heavily on the special choice of 7 for the parameter D.

A *theoretical method* is capable of proving results that hold for some large set of values of the parameters. The Gelfond-Baker theory is of a theoretical nature, since it yields upper bounds for the solutions of many equations in terms of their parameters. Other examples are application of the theory of quadratic reciprocity, that shows that $x^2 + D = 2^n$ has no solutions at all if D is odd, at least 5, and not congruent to $7 \pmod 8$, and application of the theory of hypergeometric functions, which Beukers [1981] used to show that the solutions (x,n) of $x^2 + D = 2^n$ satisfy $n < 435 + 10\cdot{}^2\log|D|$, and if $|D| < 2^{96}$ then moreover $n < 18 + 2\cdot{}^2\log|D|$. Theoretical methods are often too general to be able to produce all the solutions of a given equation.

An *algorithmic method* is a method that is guaranteed to work for any set of values of the parameters, but has to be applied separately to each particular set of parameter values, in order to produce all the solutions. The methods used in this book are mainly of such an algorithmic nature. For the equation $x^2 + D = 2^n$ (and actually for a more general equation) we will give an algorithmic method in Chapter 4. In fact, since Beukers' above-mentioned result provides a small upper bound for the solutions, it can be made algorithmic by providing a simple method of enumerating all the solutions below the upper bound. However, the algorithmic part of this method is trivial, and therefore we still prefer to classify Beukers' method as theoretical. In order to make the Gelfond-Baker theory algorithmic, enumeration of all possibilities is impractical. Therefore more ingenious ways of determining all the solutions below a large upper bound have to be found. We remark that Beukers' method for the more general equation $x^2 + D = p^n$ also has an ad hoc aspect, since it works for some special values of p only. Our method of Chapter 4 does not have this disadvantage.

An ideal towards which one might strive in solving diophantine equations is to devise a computer algorithm, a kind of 'diophantine machine', which only has to be fed with the parameters of the equation, and after a short time gives as output a list of all the solutions. One should have a guarantee (in the strictest mathematical sense of proof) that no solutions are missing.

At first sight the method outlined above, and described in this monograph, seems to be a good candidate to be developed into such a general applicable algorithm. Namely, the second step is of a quite general nature, providing upper bounds for exponential diophantine equations that are explicit in the parameters of the equation. Also the third step, the algorithmic diophantine approximation part, works in principle for any set of values substituted for the parameters. However, the computations have to be performed separately for each particular set of values.

The main difficulties in devising such a 'diophantine machine' are in the first part of the method outlined above, especially if some algebraic number theory is used. Developments taking place in the theory of algorithmic algebraic number theory on computing fundamental units and on finding factorizations of prime numbers in algebraic extensions, are of importance here. We believe that when suitable algorithms of this kind are available, it will be possible in principle to make such a 'diophantine machine' (but technical difficulties in the third step should not be underestimated). The generality of such an algorithm is restricted by the generality of the first step, the transformation to the linear form in logarithms. In this book we use computer algorithms only if the magnitude of the computational tasks makes this necessary, and keep to "manual" work otherwise. In this way we also try to keep the presentation of the methods lucid.

The reader should be aware of the fact that the computer programs and their results are part of the proofs of many of our theorems on specific diophantine equations. It is however impossible to publish all details of these programs and computations. The interested reader may obtain the details from the author by request, and is invited to check the computations himself.

The book by Shorey and Tijdeman [1986] gives a good survey of the diophantine equations for which computable upper bounds for the solutions can be found using the Gelfond-Baker method (see also Shorey, van der Poorten, Tijdeman and Schinzel [1977], and Stroeker and Tijdeman [1982]). Some of these equations can be completely solved by the methods described in this book, among which there are purely exponential equations, equations involving binary recurrence sequences, and Thue equations and Thue-Mahler equations. Especially the latter two are of importance in various other parts of number theory. For example, they are the key to solving Mordell equations and various equations arising in algebraic number theory and arithmetic algebraic geometry. The Gelfond-Baker method was used to actually solve a diophantine equation for the first time in the work of Baker and Davenport [1969] in solving the system of diophantine equations

$$3\cdot x^2 - 2 = y^2 , \quad 8\cdot x^2 - 7 = z^2 .$$

Other equations occuring in the literature for which upper bounds for the solutions can be computed, cannot be treated as easily by our algorithmic methods, because the application of the theory of linear forms in logarithms is more complicated for these equations, and moreover the upper bounds are essentially too large. An example of this kind is the Catalan equation $a^x - b^y = 1$ in integers a, b, x, y, all $\geqslant 2$. Catalan conjectured in 1844 that this equation has only the solution $(a,b,x,y) = (3,2,2,3)$. Tijdeman [1976] proved that the solutions of the Catalan equation are bounded by a computable number. This number can be taken to be $\exp(\exp(\exp(\exp(730))))$, according to Langevin [1976]. However, we fail to see how the methods that we describe in the forthcoming chapters can be applied for completely solving the Catalan equation, and we believe that Grosswald's remarks on this topic are too optimistic (Grosswald [1984], p. 259, in particular the footnote).

Another diophantine equation, that for centuries has attracted the attention of many mathematicians, is the Fermat equation $x^n + y^n = z^n$ in integers x, y, z, n, with $n \geqslant 3$ and $x\cdot y\cdot z \neq 0$. It is conjectured to have no solutions. Faltings [1983] proved that for fixed n the number of solutions is finite. His proof is ineffective. The Gelfond-Baker theory seems not to be strong enough to deal with the Fermat equation in its full generality, not even if n is fixed. For a survey of partial results on the Fermat equation that have been obtained using this theory, see Tijdeman [1985] and Chapter 11 of Shorey and Tijdeman [1986].

We remark that for many diophantine equations recently important progress has been made in determining upper bounds for the *number* of solutions. See e.g. Evertse [1983], Evertse, Györy, Stewart and Tijdeman [1988] and Schmidt [1988] for a survey. These results are often remarkably sharp, but ineffective, so that they cannot be used for actually finding the solutions.

To conclude this section we give an overview of the contents of this monograph. It is divided into three parts: Chapter 1 is introductory, Chapters 2 and 3 give the necessary preliminaries, and Chapters 4 to 8 deal with various types of diophantine equations.

Sections 1.2 to 1.5 give a short introduction for the non-specialist to respectively the Gelfond-Baker theory, diophantine approximation theory, the algorithmic aspects of diophantine approximation, and the procedure for reducing upper bounds. Chapter 2 contains the preliminary results that we need from algebraic number theory and from the theory of p-adic numbers and functions, and quotes in full detail the theorems from the Gelfond-Baker theory which we use. It concludes with some remarks on numerical methods. Chapter 3 gives in detail the algorithms in the field of diophantine approximation theory that we apply in the subsequent chapters. In a sense this chapter is the heart of the book.

Chapters 4 to 8 are each devoted to a certain type of diophantine equation. Let $p_1, \ldots, p_s$ be a fixed set of distinct primes. Let S be the set of positive integers composed of primes $p_1, \ldots, p_s$ only.

Chapter 4 deals with elements of binary recurrence sequences ("generalized Fibonacci sequences") that are in S, and gives applications to mixed quadratic-exponential equations, such as the generalized Ramanujan-Nagell equation $x^2 + k \in S$ (k fixed). The diophantine approximation part of this chapter is interesting for two reasons: the p-adic approximation is very simple, and in the case of the recurrence having negative discriminant, a nice interplay of p-adic and real/complex approximation arguments occurs. The research for Chapter 4 was done partly in cooperation with A. Pethö from Debrecen. The results have been published in Pethö and de Weger [1986] and de Weger [1986[b]].

Chapter 5 deals with the diophantine inequality $0 < x - y < y^{\delta}$, where $x, y \in S$, and $\delta \in (0,1)$ is fixed. Chapter 6 deals with $x + y = z$, where $x, y, z \in S$, which can be considered as the p-adic analogue of the inequality of Chapter 5. These two equations are the simplest examples of diophantine equations that can be treated by our method. Since they are already purely exponential equations of the form (1.1) or (1.2) with $t = 2$, the first step is trivial: the linear forms in logarithms are directly related to the equations. Therefore they serve as good examples to get a clear understanding of the diophantine approximation part of our method. The results of these chapters have been published in de Weger [1987].

Chapter 7 studies the equation $x + y = z^2$, where $x, y \in S$, and $z \in \mathbb{Z}$. This equation is a further generalization of the generalized Ramanujan-Nagell equation, studied in Chapter 4.

In Chapter 8 a procedure is given to solve Thue equations, that works in principle for Thue equations of any degree. It is applied to find all integral points on the elliptic curve $y^2 = x^3 - 4\cdot x + 1$. We also mention briefly how Thue-Mahler equations can be dealt with. This chapter has been written jointly with N. Tzanakis from Iraklion. The results have been published in Tzanakis and de Weger [1989[a]], and in de Weger [1989[a]].

1.2. The Gelfond-Baker method.

In Section 1.1 we have explained that before applying the Gelfond-Baker method to some diophantine equation, the equation should be transformed into a purely exponential diophantine equation or inequality with not too many terms (cf. (1.1), (1.2)). In this section we sketch the arguments from the Gelfond-Baker theory that lead to upper bounds for the variables of this exponential equation/inequality.

Let us first treat the case of the inequality (1.2). Since $t = 2$ we may assume that it has the form

$$\left| \alpha_0 \cdot \prod_{i=1}^{s} \alpha_i^{n_i} - 1 \right| < C_0 \cdot \exp(-\delta \cdot N) ,$$

where the α_i are fixed algebraic numbers, $N = \max|n_i|$, and C_0, δ are positive constants. In the examples we study, we encounter one of the following two cases: either all α_i are real, or $|\alpha_i| = 1$ for all i . In the real case, if N is large enough, the linear form in logarithms

$$\Lambda = \log|\alpha_0| + \sum_{i=1}^{s} n_i \cdot \log|\alpha_i|$$

must satisfy

$$|\Lambda| < C_0' \cdot \exp(-\delta \cdot N) \tag{1.3}$$

for some C_0' . In the complex case, the same inequality (1.3) follows for the linear form

$$\begin{aligned}\Lambda &= \mathrm{Log}\ \alpha_0 + \sum_{i=1}^{s} n_i \cdot \mathrm{Log}\ \alpha_i + k \cdot \mathrm{Log}(-1) \\ &= i \cdot \left(\mathrm{Arg}\ \alpha_0 + \sum_{i=1}^{s} n_i \cdot \mathrm{Arg}\ \alpha_i + k \cdot \pi \right) ,\end{aligned}$$

where the Log and Arg functions take their principal values. Now we can apply one of the many results from the Gelfond-Baker theory, giving an explicit lower bound for $|\Lambda|$ in terms of N , e.g. the following theorem.

<u>THEOREM 1.1. (Baker [1972]).</u> *Let Λ be as above. There exist computable constants C_1, C_2 , depending on the α_i only, such that if $\Lambda \neq 0$ then*

$$|\Lambda| > \exp\left(-(C_1 + C_2 \cdot \log N)\right) .$$

We usually know that $\Lambda \neq 0$. Combining (1.3) and Theorem 1.1 we then obtain

$$N < \frac{C_1 + \log C_0'}{\delta} + \frac{C_2}{\delta} \cdot \log N .$$

It follows that N is bounded from above.

Next, consider the exponential equation (1.1). By $t = 2$ we can write it as

$$\alpha_0 \cdot \prod_{i=1}^{s} \alpha_i^{n_i} - 1 = \beta_0 \cdot \prod_{j=1}^{r} \beta_j^{m_j} ,$$

where the α_i, β_j are fixed algebraic numbers. Let H_p be the maximum of the $|n_i|$, $|m_j|$ where i, j run through the set of indices for which α_i resp. β_j are non-units. Let H be the maximum of the $|n_i|$, $|m_j|$ where i, j run through the set of all indices. Suppose that p is a rational prime lying above β_j for some j . There are constants c_1, c_2 such that

$$\mathrm{ord}_p\Big(\alpha_0 \cdot \prod_{i=1}^{s} \alpha_i^{n_i} - 1\Big) \geqslant c_1 + c_2 \cdot m_j .$$

Assuming that $\mathrm{ord}_p(\alpha_i) = 0$ for all i , we may write down a p-adic linear form in logarithms

$$\Lambda = \log_p \alpha_0 + \sum_{i=1}^{s} n_i \cdot \log_p \alpha_i ,$$

for which, if m_j is large enough, it follows that

$$\mathrm{ord}_p(\Lambda) \geqslant c_1 + c_2 \cdot m_j . \tag{1.4}$$

We are now in a position to apply the following result from the p-adic Gelfond-Baker theory. Here, $N = \max|n_i|$.

THEOREM 1.2. (van der Poorten [1977], Yu [1987]). *Let* Λ , p *be as above. There exist computable constants* C_3, C_4 , *depending only on the* α_i *and on* p , *such that if* $\Lambda \neq 0$ *then*

$$\mathrm{ord}_p(\Lambda) < C_3 + C_4 \cdot \log N .$$

Applying (1.4) and Theorem 1.2 for all possible p we obtain constants C'_3, C'_4 with

$$H_p < C'_3 + C'_4 \cdot \log H .$$

If $H \leqslant C_5 \cdot H_p$ for some constant C_5 , then this immediately yields an upper bound for H . If $H > C_5 \cdot H_p$, then it can be shown that there exists a conjugate of the α_i, β_j , denoted with a prime sign, for which

$$\Big|\beta'_0 \cdot \prod_{j=1}^{r} \beta'^{\,m_j}_j\Big| < \exp(-C_6 \cdot H)$$

for a constant C_6 (cf. the proof of Theorem 1.4, pp. 45-49, of Shorey and Tijdeman [1986]). Now we can apply Theorem 1.1. This yields

$$\left|\alpha'_0 \cdot \prod_{i=1}^{s} \alpha'^{\,n_i}_i - 1\right| > \exp\left(-(C_7 + C_8 \cdot \log H)\right) .$$

It follows that H is bounded from above.

If it happens that none of the α_i, β_j are units, then of course the application of Theorem 1.2 suffices.

We remark that, in order to be able to completely solve a diophantine equation, it is crucial that all constants can be computed explicitly. Therefore we can only use the bounds from the Gelfond-Baker theory that are completely explicit. We give details of such theorems in Section 2.4.

1.3. Theoretical diophantine approximation.

In this section we briefly mention some results from diophantine approximation theory, thus giving a background to the next section. We refer to Koksma [1937], Cassels [1957] (Chapters I and III) and to Hardy and Wright [1979] (Chapters XI and XXIII), for further details.

The simplest form of diophantine approximation in the real case is that of approximation of a real number ϑ by rational numbers p/q . It is well known that if ϑ is irrational, then there exist infinitely many solutions $(p,q) \in \mathbb{Z}\times\mathbb{N}$ with $(p,q) = 1$ of the diophantine inequality

$$\left| \vartheta - \frac{p}{q} \right| < q^{-2} .$$

All convergents from the continued fraction expansion of ϑ are such solutions. The convergents are simple to compute for any particular $\vartheta \in \mathbb{R}$.

One way of generalizing this is to study simultaneous approximations to a set of real numbers $\vartheta_1, \ldots, \vartheta_n$, i.e. rational approximations to ϑ_i all having the same denominator. It is well known that the system of inequalities

$$\left| \vartheta_i - \frac{p_i}{q} \right| < q^{-(1+1/n)} \quad \text{for} \quad i = 1, \ldots, n$$

has infinitely many solutions $(p_1,\ldots,p_n,q)$ if at least one of the ϑ_i is irrational. But it is much harder to find solutions of such inequalities than in the case $n = 1$. Some multi-dimensional continued fraction algorithms have been devised (cf. Brentjes [1981] for a survey), but they seem not to have the desired simplicity and generality. We shall see later how we can apply the so-called L^3-algorithm to this problem.

Another way of generalizing the simplest case of diophantine approximation is to study linear forms, such as

$$L = \sum_{j=1}^{m} q_j \cdot \vartheta_j ,$$

where $\vartheta_1, \ldots, \vartheta_m$ are given real numbers, and $q_1, \ldots, q_m$ are the unknowns in $\mathbb{Z}$. Put $Q = \max|q_i|$. A classical theorem guarantees the existence of a solution $(p, q_1, \ldots, q_m)$ of the inequality

$$| L - p | < Q^{-m} .$$

Note that the case $m = 1$ becomes our first inequality on dividing by $q = q_1$. Also in this case the L^3-algorithm is very useful, as we shall see below.

We can incorporate the two generalizations above in a further generalization, that of simultaneous approximation of linear forms. Let real numbers ϑ_{ij} be given for $i = 1, \ldots, n$, $j = 1, \ldots, m$. Put

$$L_i = \sum_{j=1}^{m} q_j \cdot \vartheta_{ij} \quad \text{for } i = 1, \ldots, n .$$

A celebrated theorem of Minkowski states that there exists a solution $(p_1, \ldots, p_n, q_1, \ldots, q_m)$ of the system of inequalities

$$| L_i - p_i | < Q^{-m/n} \quad \text{for } i = 1, \ldots, n .$$

As we shall show in Section 1.4, the L^3-algorithm may be applied to this general form. We actually compute solutions of systems of inequalities that are slightly weaker in the sense that the right hand side is multiplied by a small constant larger than 1.

We now consider inhomogeneous approximation. This means that for all i there is an inhomogeneous term β_i in the linear form L_i, viz.

$$L_i = \beta_i + \sum_{j=1}^{m} q_j \cdot \vartheta_{ij} \quad \text{for } i = 1, \ldots, n .$$

Again, there exists a constant c such that the system

$$| L_i - p_i | < c\cdot Q^{-m/n} \quad \text{for} \quad i = 1, \ldots, n ,$$

under some independence condition on the β_i and ϑ_{ij} , has a solution. This is Kronecker's theorem. The simplest case $m = n = 1$ comes down to

$$| q\cdot\vartheta - p + \beta | < c\cdot q^{-1} .$$

The upper bounds given above, that tell us that the order of magnitude of $| L_i - p_i |$ can be at least as small as $Q^{-m/n}$, are not only theoretical upper bounds, but they predict the heuristically expected order of magnitude as well. By this we mean that in a generic situation (i.e. when there are no almost-linear relations between the ϑ_{ij} (and the β_i), it is indeed the case that for a given Q_0 the minimal $\max_i |L_i - p_i|$, taken over all $Q \leqslant Q_0$, has the order of magnitude of the upper bound $Q^{-m/n}$.

To conclude this section, we remark that there is a p-adic analogue of this theory of diophantine approximation, founded by Mahler and Lutz. If we replace in the above considerations $\mathbb{R}$ by $\mathbb{Q}_p$, the absolute value $|\cdot|$ by the p-adic value $|\cdot|_p$, and the measure Q for an approximation $(p_1,\ldots,p_n,q_1,\ldots,q_m)$ by any convex norm $\Phi(p_1,\ldots,p_n,q_1,\ldots,q_m)$ on $\mathbb{R}^{n+m}$, then the p-adic analogues of the theorems of Minkowski and Kronecker are essentially analogous to the above mentioned results in the real case. See Koksma [1937] for references to Mahler's work, and Lutz [1951], and for a detailed analysis of the case $n = 1$, $m = 2$ see de Weger [1986[a]].

1.4. Computational diophantine approximation.

In this section we give some idea of practically solving the diophantine approximation problems that we encounter in solving diophantine equations. In this section we give no rigorous treatment. We neglect worst cases, and concentrate on how things are expected to work (according to the heuristics of Section 1.3), and appear to work in practice. In the subsequent chapters many examples are given, showing that our methods are indeed useful in practice. Applying the method in practice may be the best way of acquiring the necessary *Fingerspitzengefühl* for the method.

We shall deal with the following computational diophantine approximation problem. Let ϑ_{ij}, $\beta_i \in \mathbb{R}$ be given, and let $p_1, \ldots, p_n$, $q_1, \ldots, q_m$ be integral unknowns with $Q = \max|q_j|$. Let L_i be as above. Let a positive constant Q_0, assumed to be a rather large number, 10^{50} say, be given. Find a lower bound for the value of

$$\max_i | L_i - p_i | ,$$

where $(p_1,\ldots,p_n,q_1,\ldots,q_m)$ runs through the set of values with $Q \leqslant Q_0$. From the heuristics outlined in Section 1.3 it follows that one will be satisfied if this lower bound is of the size $Q_0^{-m/n}$. For the p-adic case an analogous problem may be formulated.

Related problems in diophantine approximation theory are those of actually finding a good or the best solution of $\max_i|L_i-p_i| < \varepsilon$ for a fixed $\varepsilon > 0$. As we shall see, the L^3-algorithm is a very useful tool for finding good solutions. The problem of finding the best solution however seems to be essentially more difficult. We note that in most of our applications of solving diophantine equations it suffices to have a suitable lower bound for $\max_i|L_i-p_i|$ for a given Q_0, while it is unnecessary to know explicitly how sharp this bound is.

The computational tool that we use to solve the afore-mentioned problems is the so-called L^3-lattice basis reduction algorithm, described in Lenstra, Lenstra and Lovász [1982]. We shall give details of this algorithm in Sections 3.4 and 3.5. Below we briefly indicate how it can be used to solve diophantine approximation problems.

Let Γ be a lattice in $\mathbb{R}^n$. The L^3-algorithm accepts as input an arbitrary basis $\underline{b}_1, \ldots, \underline{b}_n$ of Γ. As output it gives another basis $\underline{c}_1, \ldots, \underline{c}_n$ of the same lattice Γ, that is a so-called *reduced* basis. The concept *reduced* means something like nearly orthogonal. From a reduced basis it is possible to compute lower bounds for the following two quantities:

→ the length of the non-zero lattice point that is nearest to the origin:

$$\ell(\Gamma) = \min_{\underline{0}\neq\underline{x}\in\Gamma} |\underline{x}| ,$$

(see Lenstra, Lenstra and Lovász [1982], Prop. (1.11), and our Lemma 3.4),

$\rightarrow$ for any given point $\underline{y} \in \mathbb{R}^n$, the distance from $\underline{y}$ to the nearest lattice point:

$$\ell(\Gamma,\underline{y}) = \min_{\underline{x}\in\Gamma} |\underline{x}-\underline{y}| ,$$

(see Babai [1986], and our Lemmas 3.5 and 3.6).

The L^3-algorithm enjoys the property that these lower bounds are usually near to the actual minimal solutions. In a generic situation, where the lattice is not too distorted, the vectors $\underline{c}_i$ of the reduced basis all have about the same length, which is of the order of magnitude of

$$\det(\Gamma)^{1/n} .$$

The value of $\ell(\Gamma)$ as well as the lower bounds computed for it, are about as large as that. If $\underline{y}$ is not too close to a lattice point, the same holds for $\ell(\Gamma,\underline{y})$. Moreover, the running time of the algorithm is good, both in the theoretical sense (it is polynomial-time in the length of the input-parameters), and in practice (cf. Lenstra [1984], p. 7).

To solve the problem of finding a lower bounds for $\max_i |L_i-p_i|$ as formulated above, we take the lattice Γ as follows. Let C be an integer, at least as large as $Q_0^{1+m/n}$. The lattice Γ , of dimension $n + m$, is defined by specifying a basis, namely the column vectors $\underline{b}_1, \ldots, \underline{b}_{n+m}$ of the matrix

$$\mathcal{B} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & \emptyset & \\ \emptyset & & 1 & & & \\ [C\cdot\vartheta_{11}] & \ldots & [C\cdot\vartheta_{1m}] & -C & & \\ \vdots & & \vdots & & \ddots & \\ [C\cdot\vartheta_{n1}] & \ldots & [C\cdot\vartheta_{nm}] & \emptyset & & -C \end{pmatrix} .$$

(The symbol $\emptyset$ means that all not explicitly given entries in that area are zero). Applying the L^3-algorithm to this lattice we find a reduced basis, of which the basis vectors will have lengths of about $C^{n/(m+n)}$, which is roughly the size of Q_0 . Generally speaking, the larger C is, the larger the lengths of the basis vectors of a reduced basis will be (and the larger the lower bounds for $\ell(\Gamma)$ and $\ell(\Gamma,\underline{y})$ will be).

Let us first treat the homogeneous case, i.e. $\beta_i = 0$ for all i . Consider the lattice point $\underline{x} = \mathcal{B}\cdot(q_1,\ldots,q_m,p_1,\ldots p_n)^T$. It is equal to

$$\underline{x} = (q_1,\ldots,q_m,\tilde{L}_1-C\cdot p_1,\ldots,\tilde{L}_n-C\cdot p_n)^T ,$$

where

$$\tilde{L}_i = \sum_{j=1}^{m} q_j\cdot[C\cdot\vartheta_{ij}] \quad \text{for} \quad i = 1, \ldots, n .$$

From the application of the L^3-algorithm we find a lower bound for $\ell(\Gamma)$, of size Q_0 . We assume it to be large enough (if this is not the case, we try a somewhat larger value for C , and perform the L^3-algorithm again for the lattice defined for this C). So we may assume that there is a small constant c_1 such that

$$\sum_{i=1}^{n} (\tilde{L}_i-C\cdot p_i)^2 \geqslant \ell(\Gamma)^2 - m\cdot Q_0^2 > c_1\cdot Q_0^2 .$$

We have $|\tilde{L}_i-C\cdot L_i| \leqslant m\cdot Q_0$, so we may assume that for small constants c_2, c_3

$$\max_i|L_i-p_i| > c_2\cdot C^{-1}\cdot\max|\tilde{L}_i-C\cdot p_i| > c_3\cdot Q_0/C .$$

By the choice of C this last bound has the required size.

Next, we study the inhomogeneous case, where not all β_i are zero. We take the same lattice Γ as in the homogeneous case (note that the lattice definition depends only on the ϑ_{ij} and the C). Consider the point

$$\underline{y} = (0,\ldots,0,-[C\cdot\beta_1],\ldots,-[C\cdot\beta_n])^T .$$

From the reduced basis found by the L^3-algorithm we have a lower bound for $\ell(\Gamma,\underline{y})$. Assume that it is large enough, and of size Q_0 . We take the same lattice point $\underline{x} = \mathcal{B}\cdot(q_1,\ldots,q_m,p_1,\ldots p_n)^T$ as in the homogeneous case. Then

$$\underline{x} - \underline{y} = (q_1,\ldots,q_m,\tilde{L}_1-C\cdot p_1,\ldots,\tilde{L}_n-C\cdot p_n)^T ,$$

where

$$\tilde{L}_i = [C\cdot\beta_i] + \sum_{j=1}^{m} q_j\cdot[C\cdot\vartheta_{ij}] \quad \text{for} \quad i = 1, \ldots, n .$$

The same reasoning as in the homogeneous case now yields the desired result. Note that if we have performed the L^3-algorithm once for given ϑ_{ij} , we may use the result to treat the homogeneous case, and many inhomogeneous cases with different β_i 's as well, as long as the ϑ_{ij}'s are the same.

The above process describes how to find lower bounds for systems of diophantine inequalities. It will be clear from the above that it is not difficult to find good solutions, i.e. $(q_1,\ldots,q_m, p_1,\ldots,p_n)$ with $Q \leqslant Q_0$ and $\max_i |L_i - p_i|$ near to the best possible value. In particular, the basis vectors of a reduced basis are adequate for the homogeneous case, and for the inhomogeneous case the lattice points near to $\underline{y}$ will be such solutions. The lattice points near to $\underline{y}$ are not difficult to find once a reduced basis is available. Specifically, if $s_1, \ldots, s_n \in \mathbb{R}$ are the coordinates of $\underline{y}$ with respect to a reduced basis, then one may take the lattice points with coordinates (with respect to the reduced basis) $t_i \in \mathbb{Z}$ that are near to s_i for $i = 1, \ldots, n$.

In the definition of the matrix above the expressions $[C\cdot\vartheta_{ij}]$ occur. Using these expressions we have constructed a lattice Γ that is completely integral, i.e. $\Gamma \subset \mathbb{Z}^{m+n}$. The L^3-algorithm can be adapted to work exact for those lattices, so that rounding-off errors are avoided (cf. Section 3.5). The "errors" occur only in the difference between the $\tilde{L}_i$ and the $C\cdot L_i$, and are thus kept under control by choosing the proper constants c_1, c_2, c_3 . Of course one should take care to have the numerical values of the ϑ_{ij} and the β_i correct to sufficient precision. We shall discuss such numerical problems briefly in Section 2.5.

A possible variation of the above diophantine approximation problem is to give weights to the linear forms L_i , i.e. to look for a lower bound for

$$\max_i w_i \cdot |\ L_i - p_i\ | ,$$

where the w_i are fixed positive numbers. This situation can be dealt with easily by replacing every C in the $(n+i)$ th row of the matrix by $C\cdot w_i$.

Another variation is the problem where not all the variables q_j have the same upper bound Q_0 . To illustrate this, assume that $n = 1$, and that

$$L = \sum_{j=1}^{m} q_j\cdot\vartheta_j .$$

Now suppose that for some $Q_1 > Q_2$ (it will be handy to have $Q_2 \mid Q_1$) we are interested in the solutions with

$$|q_j| \leqslant Q_1 \quad \text{for} \quad j \leqslant m_1 , \quad |q_j| \leqslant Q_2 \quad \text{for} \quad j \geqslant m_1+1 .$$

Next, let C be of the size of $Q_1^{m_1+1} \cdot Q_2^{m-m_1}$, and take the matrix

$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & \emptyset & & \\ & & 1 & & & & & \\ & \emptyset & & Q_1/Q_2 & & & & \\ & & & & \ddots & & & \\ & & & & & Q_1/Q_2 & & \\ [C\cdot\vartheta_1] & \dots & [C\cdot\vartheta_{m_1}] & [C\cdot\vartheta_{m_1+1}] & \dots & [C\cdot\vartheta_m] & -C \end{pmatrix}.$$

Its determinant is of the size of Q_1^{m+1}. For a lattice point $(q_1,\dots,q_m,\tilde{L}-C\cdot p)^T$ we therefore expect that $\max(|q_1|,\dots,|q_{m_1}|)$, $(Q_1/Q_2)\cdot\max(|q_{m_1+1}|,\dots,|q_m|)$ and $|\tilde{L}-C\cdot p|$ are all of the size of Q_1. It follows that $|L-p|$ is of the size of $Q_1^{-m_1}\cdot Q_2^{-(m-m_1)}$, in accordance with the heuristics. This variant is useful when a combination of real and p-adic techniques is used, such as for the Thue-Mahler equation (see Section 8.6).

We conclude this section by giving the analogous method of p-adic diophantine approximation. We assume that the ϑ_{ij}, β_i are in $\mathbb{Q}_p$, and, moreover, that they are p-adic integers. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For any p-adic integer γ and any $\mu \in \mathbb{N}_0$ we denote by $\gamma^{(\mu)}$ the unique rational integer such that

$$\gamma \equiv \gamma^{(\mu)} \pmod{p^\mu}, \quad 0 \leqslant \gamma^{(\mu)} < p^\mu .$$

Let $\mu \in \mathbb{N}$ be such that p^μ is roughly the same size as $Q_0^{1+m/n}$, and assume that μ is large enough (it is the analogue of the constant C in the real case above). Take for Γ the lattice of which a basis is given by the column vectors of the matrix

$$\mathcal{B} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & \emptyset & \\ \emptyset & & 1 & & & \\ \vartheta_{11}^{(\mu)} & \dots & \vartheta_{1m}^{(\mu)} & p^\mu & & \\ \vdots & & \vdots & & \ddots & \\ \vartheta_{n1}^{(\mu)} & \dots & \vartheta_{nm}^{(\mu)} & \emptyset & & p^\mu \end{pmatrix}.$$

Consider the lattice point

$$\mathcal{B}\cdot(q_1,\dots,q_m,z_1,\dots,z_n)^T = (q_1,\dots,q_m,p_1,\dots,p_n)^T .$$

Then it is obvious that

$$p_i = \sum_{j=1}^{m} q_j \cdot \vartheta_{ij}^{(\mu)} + z_i \cdot p^{\mu} .$$

Hence the lattice Γ can be described as the set

$$\Gamma = \langle \, (q_1,\ldots,q_m,p_1,\ldots,p_n)^T \in \mathbb{Z}^{m+n} \mid \sum_{j=1}^{m} q_j \cdot \vartheta_{ij} \equiv p_i \pmod{p^{\mu}} \quad \text{for} \quad i = 1, \ldots, n \, \rangle .$$

The L^3-algorithm provides a lower bound for the length of the nonzero vectors in this set, which is of the same size as $p^{\mu \cdot n/(n+m)}$, and that of Q_0 . This yields the desired result, if μ is taken large enough.

For the inhomogeneous case, put

$$\underline{y} = (0,\ldots,0,-\beta_1^{(\mu)},\ldots,-\beta_n^{(\mu)})^T ,$$

and consider the set

$$\Gamma^* = \langle \, (q_1,\ldots,q_m,p_1,\ldots,p_n)^T \in \mathbb{Z}^{m+n} \mid \beta_i + \sum_{j=1}^{m} q_j \cdot \vartheta_{ij} \equiv p_i \pmod{p^{\mu}} \quad \text{for} \quad i = 1, \ldots, n \, \rangle .$$

Then $\underline{x} \in \Gamma^*$ if and only if $\underline{x} + \underline{y} \in \Gamma$, so Γ^* is a translated lattice. A lower bound for $\ell(\Gamma,\underline{y})$ now yields the desired result.

Again variations are possible, as in the real case, e.g. by replacing on the $(n+i)$ th row the μ by different μ_i . It is even possible in this way to treat more than one prime p at the same time, by replacing on the $(n+i)$ th row the p^{μ} by different $p_i^{\mu_i}$.

We indicate one more variation for the p-adic case. Suppose we have only one linear form $\Lambda = \sum_{j=1}^{m} q_j \cdot \vartheta_j$, and one variable $p \in \mathbb{Z}$, and we want to study when Λ is congruent to 0 modulo different prime powers $p_1^{\mu_1}, \ldots, p_n^{\mu_n}$. Thus we are interested in the set

$$\Gamma' = \langle \, [\![q_1,\ldots,q_m,p]\!]^T \in \mathbb{Z}^{m+1} \mid \sum_{j=1}^{m} q_j \cdot \vartheta_j \equiv p \pmod{p_i^{\mu_i}} \quad \text{for} \quad i = 1, \ldots, n \, \rangle$$

Then we take $\vartheta_j^* \in \mathbb{Z}$ with

$$\vartheta_j^* \equiv \vartheta_j \pmod{p_i^{\mu_i}} \quad \text{for} \quad i = 1, \ldots, n \ , \quad 0 \leqslant \vartheta_j^* < \prod_{i=1}^{n} p_i^{\mu_i} \ ,$$

for all j. The ϑ_j^* can be computed by the Chinese Remainder Theorem. Now Γ' is the lattice generated by the column vectors of

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \emptyset \\ & \emptyset & 1 & \\ \vartheta_1^* & \ldots & \vartheta_m^* & \prod_{i=1}^{n} p_i^{\mu_i} \end{pmatrix} ,$$

and we proceed with this lattice as described above.

We conclude this section with three remarks. Firstly, in the case that the dimension of the lattice under consideration is only 2, the L^3-algorithm is essentially the continued fraction algorithm, and so yields nothing new. For the p-adic continued fraction algorithm, see de Weger [1986[a]]. Secondly, the inhomogeneous case of diophantine approximation of one linear form of real numbers can also be treated by what is known as Davenport's lemma, cf. Baker and Davenport [1969] (and its multi-dimensional generalization, cf. Ellison [1971[a]]). We will return to this in Chapter 3, and explain there why we prefer our method.

Finally, one of the nice features of the above method of practical diophantine approximation is that if an extreme solution exists, then in the homogeneous case the lattice (with proper constant C or μ) will be distorted. This means that the reduced basis will not be as nice as expected, for example there might be a basis vector in it that is substantially shorter than the other ones. In the inhomogeneous case the existence of an extreme solution means that there is a lattice point extremely near to $\underline{y}$. The algorithm detects such an extraordinary situation at once, and in most cases the extremal solution is presented explicitly (e.g. in the homogeneous case as one of the vectors of the reduced basis). One can check whether this extremal solution actually satisfies the original equation, and then proceed by replacing in the above reasoning $\ell(\Gamma)$ or $\ell(\Gamma,\underline{y})$ by lower bounds for all vectors in the lattice except the extremal one. These new lower bounds will in general be of the expected size. However, when we solved diophantine equations in practice, we have never met such an extraordinary situation.

1.5. The procedure for reducing upper bounds.

We have seen in Section 1.2 how upper bounds for the solutions of the exponential inequalities and equations occuring there can be found. In Section 1.4 we have studied some diophantine approximation theory from a practical point of view. Now these two things come together.

From the application of the Gelfond-Baker theory we are left with the following problem. We have a linear form

$$\Lambda = \beta + \sum_{j=1}^{m} n_j \cdot \vartheta_j ,$$

where the β and ϑ_j are constants (that they are logarithms of algebraic numbers is now of no importance anymore), and the n_j are integral unknowns. We know that Λ is extremely close to 0, namely

$$|\Lambda| < c \cdot \exp(-\delta \cdot N) ,$$

where c, δ are (small) constants, and $N = \max|n_j|$. Finally, we have an explicit upper bound N_0 for N. This N_0 is very large, 10^{50} say.

It will be clear from Section 1.4 that the methods outlined there are of use for solving this problem. For Q_0 we take N_0. We have $n = 1$. In the real case we expect, by choosing C at least of size N_0^{m+1}, that

$$|\Lambda| > c' \cdot N_0^{-m} ,$$

for a small constant c'. It follows by combining the two inequalities for $|\Lambda|$ that

$$N < \log(c/c')/\delta + (m/\delta) \cdot \log N_0 .$$

So the upper bound N_0 for N is reduced to an upper bound N_1 of the size of $\log N_0$, which is a considerable improvement indeed. We now may apply the procedure with N_1 instead of N_0, and repeat until no further improvement is obtained. In practice it appears almost always to be the case that in that situation the reduced upper bound is near to the actual largest solution, anyway so small that simple methods of finding all the solutions below that bound suffice.

In the p-adic case an analogous reduction of upper bounds can be reached, following a similar argument. We have for the linear form Λ (cf. (1.4)),

$$\mathrm{ord}_p(\Lambda) \geqslant c_1 + c_2 \cdot m_j ,$$

where c_1, c_2 are small constants, and m_j is one of the variables. Moreover, the variables are bounded by a large constant N_0 , that is explicitly known. We take μ such that p^μ is at least of size N_0^{m+1} , so that the lower bound for the shortest nonzero vector in Γ (or Γ^*) is larger than $\sqrt{m} \cdot N_0$. Then it follows that the elements of the lattice Γ (or of the translated lattice Γ^*) cannot be solutions of (1.2). Therefore,

$$c_1 + c_2 \cdot m_j < \mu ,$$

so that we find a new upper bound for m_j , that is of the size of μ , which is about $\log N_0 / \log p$. We repeat this procedure for all the m_j , in order to obtain a reduced upper bound for H_p . If this is not yet sufficient to derive at once a reduced upper bound for H , then we can do so by applying a reduction step for real linear forms, where we may take advantage of the fact that for some of the variables a much better upper bound has just been found (cf. the second variation in Section 1.4). Again we repeat the whole procedure as far as possible.

Chapter 2. Preliminaries.

2.1. Algebraic number theory.

In this section we quote results from algebraic number theory that we use throughout the remaining chapters. We refer to Borevich and Shafarevich [1966] or any other textbook on algebraic number theory for full details.

Let K be a finite algebraic extension of $\mathbb{Q}$, of degree $D = [K:\mathbb{Q}]$. There are D embeddings $\sigma : K \to \mathbb{C}$. Let $\alpha \in K$ be an element of degree d, and let $a_0 > 0$ be the leading coefficient of its minimal polynomial over $\mathbb{Z}$. We define the *(logarithmic) height* $h(\alpha)$ by

$$h(\alpha) = \frac{1}{D}\cdot\log\Big(a_0^{D/d}\cdot\prod_{\sigma}\max(1,|\sigma(\alpha)|)\Big) ,$$

where the product is taken over all embeddings σ. Note that this definition does not depend on the field K. Hence, if the conjugates of α are $\alpha = \alpha_1, .., \alpha_d$, then the above definition applied for $K = \mathbb{Q}(\alpha)$ yields

$$h(\alpha) = \frac{1}{d}\cdot\log\Big(a_0\cdot\prod_{i=1}^{d}\max(1,|\alpha_i|)\Big) .$$

In particular, if $\alpha \in \mathbb{Q}$, then with $\alpha = p/q$ for $p, q \in \mathbb{Z}$ with $(p,q) = 1$ we have $h(\alpha) = \log\max(|p|,|q|)$, and if $\alpha \in \mathbb{Z}$ then $h(\alpha) = \log|\alpha|$.

Let there be s real and $2\cdot t$ non-real embeddings (with $D = s + 2\cdot t$). Then Dirichlet's Unit Theorem states that there exists a system of $r = s + t - 1$ independent units $\varepsilon_1, ..., \varepsilon_r$, such that the group of units of K is given by

$$\langle\ \zeta\cdot\varepsilon_1^{a_1}\cdot\ldots\cdot\varepsilon_r^{a_r}\ |\ \zeta \text{ a root of unity, } a_i \in \mathbb{Z} \text{ for } i=1,\ldots,r\ \rangle .$$

There are only finitely many roots of unity in K. Any set of independent units that generate the torsion-free part of the unit group is called a system of *fundamental units*.

The number α is called an *algebraic integer* if $a_0 = 1$. Let the *norm* of an element $\alpha \in K$ be defined by

$$N_{K/\mathbb{Q}}(\alpha) = \prod_{\sigma} \sigma(\alpha) = \Big(\prod_{i=1}^{d} \alpha_i\Big)^{D/d} .$$

For algebraic integers, $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$. The units are precisely the elements of norm ± 1 . Two elements α, β of K are called *associates* if there is a unit ε such that $\alpha = \varepsilon \cdot \beta$. Let (α) denote the ideal generated by α . Associated elements generate the same ideal, and distinct generators of an ideal are associated. There exist only finitely many non-associated algebraic integers in K with given norm. The ring of algebraic integers is denoted by $\mathcal{O}_K$. Let $\alpha_1, \ldots, \alpha_D$ be elements of $\mathcal{O}_K$ that are $\mathbb{Q}$-linearly independent. Then $\mathbb{Z}\cdot\alpha_1 \times \ldots \times \mathbb{Z}\cdot\alpha_D$ is called an *order* of K if it is a subring of the 'maximal order' $\mathcal{O}_K$.

In K any algebraic integer can be written as a product of irreducible elements. Here an *irreducible* element (*prime* element) is an element that has no integral divisors but its own associates. However, this decomposition into primes need not be unique. Ideals can also be decomposed into prime ideals, and this decomposition is unique. A *principal ideal* is an ideal generated by a single element α . Two fractional ideals are called equivalent if their quotient is principal. It is well known that there are only finitely many equivalence classes. Their number is called the *class number* h_K . For an ideal $\mathfrak{a}$ it is always true that $\mathfrak{a}^{h_K}$ is a principal ideal. The norm of the (integral) ideal $\mathfrak{a}$ is defined by $N_{K/\mathbb{Q}}(\mathfrak{a}) = \#(\mathcal{O}_K/\mathfrak{a})$.

For a prime ideal $\mathfrak{p}$ there is always a rational prime number p such that $\mathfrak{p}$ is a divisor of (p) . We say that $\mathfrak{p}$ *lies above* p . The *ramification index* $e_{\mathfrak{p}}$ is the largest power to which $\mathfrak{p}$ divides (p) . The *residue class degree* $f_{\mathfrak{p}}$ is the integer such that

$$N_{K/\mathbb{Q}}(\mathfrak{p}) = p^{f_{\mathfrak{p}}} .$$

We denote by $\mathrm{ord}_{\mathfrak{p}}(\mathfrak{a})$ the exact power to which the prime ideal $\mathfrak{p}$ divides the ideal $\mathfrak{a}$. For fractional ideals $\mathfrak{a}$ this number can of course be negative. For numbers α we write $\mathrm{ord}_{\mathfrak{p}}(\alpha)$ for $\mathrm{ord}_{\mathfrak{p}}((\alpha))$. Note that

$$\mathrm{ord}_p(\alpha) = \mathrm{ord}_{\mathfrak{p}}(\alpha)/e_{\mathfrak{p}}$$

can be defined for all $\alpha \in K$. We will return to this in Section 2.3, which deals with p-adic number theory.

2.2. Some auxiliary lemmas.

In this section we give a few simple auxiliary lemmas. The first one enables us to find an upper bound in closed form for some real number $x > 1$ that is bounded by a polynomial in $\log x$. See Pethö and de Weger [1986], Lemma 2.3.

<u>LEMMA 2.1.</u> *Let* $a \geqslant 0$, $h \geqslant 1$, $b > 0$, *and let* $x \in \mathbb{R}$, $x > 1$ *satisfy*

$$x \leqslant a + b\cdot(\log x)^h .$$

If $b > (e^2/h)^h$ *then*

$$x < 2^h\cdot\left(a^{1/h}+b^{1/h}\cdot\log(h^h\cdot b)\right)^h ,$$

and if $b \leqslant (e^2/h)^h$ *then*

$$x \leqslant 2^h\cdot\left(a^{1/h}+2\cdot e^2\right)^h .$$

<u>Proof.</u> We may assume that x is the largest solution of

$$x = a + b\cdot(\log x)^h .$$

By $(z_1+z_2)^{1/h} \leqslant z_1^{1/h} + z_2^{1/h}$ we infer

$$x^{1/h} \leqslant a^{1/h} + c\cdot\log(x^{1/h}) ,$$

where $c = h\cdot b^{1/h}$. Define y by $x^{1/h} = (1+y)\cdot c\cdot\log c$. From

$$\log c < \log(c\cdot\log c)$$

it follows that

$$c^h\cdot(\log c)^h < b\cdot\left(\log\left(c^h\cdot(\log c)^h\right)\right)^h ,$$

which implies $x > c^h\cdot(\log c)^h$. Hence $y > 0$. Now,

$$\begin{aligned}(1+y)\cdot c\cdot\log c = x^{1/h} &\leqslant a^{1/h} + c\cdot\log(1+y) + c\cdot\log c + c\cdot\operatorname{loglog} c\\ &< a^{1/h} + c\cdot y + c\cdot\log c + c\cdot\operatorname{loglog} c .\end{aligned}$$

Hence

$$y\cdot c\cdot(\log c - 1) < a^{1/h} + c\cdot\operatorname{loglog} c .$$

If $c \geqslant e^2$ it follows that

$$x^{1/h} = c\cdot\log c + y\cdot c\cdot\log c < c\cdot\log c + \frac{\log c}{\log c - 1}\cdot(a^{1/h}+c\cdot\log\log c)$$
$$< 2\cdot(a^{1/h}+c\cdot\log c) .$$

If $c \leqslant e^2$, then note that $x \leqslant a + (e^2/h)^h\cdot(\log x)^h$. So we may assume $c = e^2$ in this case. The result follows. □

The next lemmas make explicit that x and $\log(1+x)$ are near if $|x|$ is small in the real and complex case, respectively.

LEMMA 2.2. *Let* $a \in \mathbb{R}$. *If* $a < 1$ *and* $|x| < a$ *then*

$$|\log(1+x)| < \frac{-\log(1-a)}{a}\cdot|x| ,$$

and

$$|x| < \frac{a}{1-e^{-a}}\cdot|e^x-1| .$$

Proof. Note that $\log(1+x)/x$ is a strictly positive and strictly decreasing function for $|x| < 1$. Hence it is for $|x| < a$ always less than its value at $x = -a$. The same is true for the function $x/(e^x-1)$. □

LEMMA 2.3. *Let* $0 < a \leqslant \pi$. *If* $|x| < a$ *then*

$$|x| < \frac{a}{2\cdot\sin(a/2)}\cdot|e^{i\cdot x}-1| .$$

If $a < 2$, $|e^{i\cdot x}-1| < a$ *and* $|x| < \pi$ *then*

$$|x| < \frac{2\cdot\arcsin(a/2)}{a}\cdot|e^{i\cdot x}-1| .$$

Proof. Note that $|e^{i\cdot x}-1| = 2\cdot|\sin(\frac{1}{2}\cdot x)|$. and that $2\cdot\sin(\frac{1}{2}\cdot x)/x$ is a positive and even function, that decreases on $0 \leqslant x < a$. Hence it takes its minimal value at $x = a$. The first inequality now follows. The second one can be proved in a similar way. □

2.3. p-adic numbers and functions.

In this section we mention the facts about p-adic numbers and functions that we use. For details we refer to Bachman [1964] and Koblitz [1977], [1980].

We assume that the reader is familiar with the field of p-adic numbers $\mathbb{Q}_p$ and the p-adic valuation ord_p . Note that the ordinary ord_p as defined in $\mathbb{Q}_p$ coincides with the definition given in Section 2.1. We denote by Ω_p the completion of the algebraic closure of $\mathbb{Q}_p$, i.e. the field to which all p-adic theory is applied.

Every nonzero number $\alpha \in \mathbb{Q}_p$ has a p-adic expansion

$$\alpha = \sum_{i=k}^{\infty} u_i \cdot p^i ,$$

where $k = \mathrm{ord}_p(\alpha)$ and the p-*adic digits* u_i are in $\{ 0, 1, \ldots, p-1 \}$, with $u_k \neq 0$. The number 0 can be represented in this way by taking $k = 0$ and all digits equal to 0 , and $\mathrm{ord}_p(0) = \infty$ by definition. If $\mathrm{ord}_p(\alpha) \geqslant 0$ then α is called a p-*adic integer*. The set of p-adic integers is denoted by $\mathbb{Z}_p$. A p-*adic unit* is an $\alpha \in \mathbb{Q}_p$ with $\mathrm{ord}_p(\alpha) = 0$. For any p-adic integer α and any $\mu \in \mathbb{N}_0$ there exists a unique rational integer $\alpha^{(\mu)} = \sum_{i=0}^{\mu-1} u_i \cdot p^i$ satisfying

$$\mathrm{ord}_p(\alpha-\alpha^{(\mu)}) \geqslant \mu , \quad 0 \leqslant \alpha^{(\mu)} \leqslant p^{\mu} - 1 .$$

For $\mathrm{ord}_p(\alpha) \geqslant k$ we also write $\alpha \equiv 0 \pmod{p^k}$. The p-*adic norm* is defined by

$$|\alpha|_p = p^{-\mathrm{ord}_p(\alpha)} .$$

In Section 2.1 we have seen how to define ord_p and ord_p on algebraic extensions of $\mathbb{Q}$. For any $\alpha \in \Omega_p$ with $\mathrm{ord}_p(\alpha) > 1/(p-1)$ we can define the p-*adic logarithm* $\log_p(1+\alpha)$ by the Taylor series

$$\log_p(1+\alpha) = \alpha - \alpha^2/2 + \alpha^3/3 - \ldots .$$

This logarithmic function has the well known properties of a logarithm, such as $\log_p(\xi_1 \cdot \xi_2) = \log_p(\xi_1) + \log_p(\xi_2)$ for all ξ_1, ξ_2 for which it is defined. Further, $\log_p(\xi) = 0$ if and only if ξ is a root of unity. In $\mathbb{Q}_p$ the only roots of unity are the $(p-1)$ th roots of unity (if p is odd). Using these properties, this logarithmic function can be extended to all $\xi \in \Omega_p$ with $\mathrm{ord}_p(\xi) = 0$, as follows. By Fermat's theorem for algebraic number fields there is a $k \in \mathbb{N}$ such that $\mathrm{ord}_p(\xi^k-1) > 1/(p-1)$. Then

$$\log_p(\xi) = \frac{1}{k}\cdot \log_p\bigl(1+(\xi^k-1)\bigr) \ .$$

An equivalent definition is $\log_p(\xi) = \log_p(\xi/\zeta)$, where ζ is a root of unity such that $\mathrm{ord}_p(\xi-\zeta) > 0$. In this way the p-adic logarithm is a well defined function. Note that $\log_p(\xi)$ lies in the subfield of Ω_p generated by ξ . Finally we note that if $\mathrm{ord}_p(\xi) > 1/(p-1)$ then

$$\mathrm{ord}_p(\xi) = \mathrm{ord}_p(\log_p(1+\xi)) \ .$$

2.4. Lower bounds for linear forms in logarithms.

In this section we quote in detail the results from the Gelfond-Baker theory that we use. They yield lower bounds for linear forms in logarithms of algebraic numbers. We do not always give the theorems in their full generality, since in this book only linear forms with rational unknowns occur, whereas most Gelfond-Baker theorems are formulated for linear forms with algebraic unknowns. We selected bounds with fully explicit constants, because only such completely explicit results can be used for our purposes.

The first result in this field for a linear form in logarithms with at least three terms is due to Baker [1966], and in the p-adic case to Coates [1969], [1970]. For a survey of this theory, see Baker [1977] and van der Poorten [1977]. We will use more recent, sharper results, due to Waldschmidt [1980] and Yu [1987]. Further improvements of the constants have been reached (see the references after Lemma 2.4 below), but too recently to be taken into account here.

First we deal with real/complex linear forms in logarithms. We quote the result of Waldschmidt [1980].

<u>LEMMA 2.4 (Waldschmidt).</u> *Let* K *be a number field with* $[K:\mathbb{Q}] = D$ *. Let* $\alpha_1, \ldots, \alpha_n \in K$ *, and* $b_1, \ldots, b_n \in \mathbb{Z}$ $(n \geqslant 2)$ *. Let* $V_1, \ldots, V_n$ *be positive real numbers satisfying* $1/D \leqslant V_1 \leqslant \ldots \leqslant V_n$ *and*

$$V_j \geqslant \max\bigl(\, h(\alpha_j),\ |\log \alpha_j|/D \,\bigr) \quad \text{for} \quad j = 1, \ldots, n \ .$$

where $\log \alpha_j$ *for* $j = 1, \ldots, n$ *is an arbitrary but fixed determination of the logarithm of* α_j *. Let* $V_j^+ = \max(V_j, 1)$ *for* $j = n, n-1$ *, and put*

$$\Lambda = \sum_{j=1}^{n} b_j \cdot \log \alpha_j \ .$$

Put $B = \max_{1 \leqslant i \leqslant n} |b_i|$ *. If* $\Lambda \neq 0$ *then*

$$|\Lambda| > \exp\Big(-2^{e(n)} \cdot n^{2\cdot n} \cdot D^{n+2} \cdot V_1 \cdot \ldots \cdot V_n \cdot \log(e \cdot D \cdot V_{n-1}^{+}) \cdot$$

$$\cdot \Big(\log B + \log(e \cdot D \cdot V_n^{+}) \Big) \Big) \ ,$$

where $e(n) = \min\big(8\cdot n + 51,\ 10\cdot n + 33,\ 9\cdot n + 39 \big)$ *. If, moreover, it is known that* $[\mathbb{Q}(\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_n}) : \mathbb{Q}] = 2^n$ *, then we can take* $e(n) = 9\cdot n + 26$ *and replace the factor* $n^{2\cdot n}$ *in the above bound for* $|\Lambda|$ *by* n^{n+4} *.*

Waldschmidt's main theorem does not give the constant $e(n)$ as detailed as we do, but he does so in his proof, cf. p. 283. We remark that improvements of the above bounds have recently been found by Blass, Glass, Manski, Meronk and Steiner [1988[a]], [1988[b]], Loxton, Mignotte, van der Poorten and Waldschmidt [1987], Philippon and Waldschmidt [1988], and Wüstholz [1988]. For the case $n = 2$, the sharpest bound has been given by Mignotte and Waldschmidt [1978], improved again by Mignotte and Waldschmidt [1988].

In the p-adic case we quote two results: one due to Schinzel [1967] (Theorem 1) for the case of a linear form in logarithms with two terms, and another for the general case, due to Yu [1987] (Theorem 1, see also Yu [1988]). We note that Yu's bounds improve much upon the results of van der Poorten [1977]. Moreover, van der Poorten's proofs seem to contain some errors. We give Schinzel's result for quadratic fields only.

<u>LEMMA 2.5 (Schinzel).</u> *Let* p *be prime. Let* Δ *be a squarefree integer, and let* D *be the discriminant of* $K = \mathbb{Q}(\sqrt{\Delta})$ *. Let* $\xi = \xi''/\xi'$ *and* $\chi = \chi''/\chi'$ *be elements of* K *, where* $\xi', \xi'', \chi', \chi''$ *are algebraic integers. Put*

$$L = \log \max\Big(|e\cdot D|^{1/4},\ \|\xi'\cdot\chi'\|,\ \|\xi'\cdot\chi''\|,\ \|\xi''\cdot\chi'\|,\ \|\xi''\cdot\chi''\| \Big) \ ,$$

where $\|\gamma\|$ *denotes the maximal absolute value of the conjugates of* $\gamma \in K$ *. Let* $\mathfrak{p}$ *be a prime ideal of* K *with norm* $N\mathfrak{p} = p^{\rho}$ *. Put* $\psi = 2/\rho \cdot \log p$ *,* $\varphi = \mathrm{ord}_{\mathfrak{p}}(p)$ *. If* ξ *or* χ *is a* $\mathfrak{p}$*-adic unit and* $\xi^n \neq \chi^m$ *, then*

$$\mathrm{ord}_{\mathfrak{p}}(\xi^n - \chi^m) < 10^6 \cdot \psi^7 \cdot \varphi^{-2} \cdot L^4 \cdot p^{4\cdot\rho+4} \cdot \big(\log \max(|m|,|n|) + \varphi \cdot L \cdot p^{\rho} + 2/L\big)^3 \ .$$

LEMMA 2.6 (Yu). *Let* $\alpha_1, \ldots, \alpha_n$ ($n \geq 2$) *be nonzero algebraic numbers. Put* $L = \mathbb{Q}(\alpha_1,\ldots,\alpha_n)$, $d = [L:\mathbb{Q}]$. *Let* $b_1, \ldots, b_n$ *be rational integers. Let* $\mathfrak{p}$ *be a prime ideal of* L *, lying above the rational prime* p *. Let* $e_{\mathfrak{p}}$ *be the ramification index, and* $f_{\mathfrak{p}}$ *the residue class degree of* $\mathfrak{p}$ *. Write* $L_{\mathfrak{p}}$ *for the completion of* L *with respect to* $\mathrm{ord}_{\mathfrak{p}}$ *(then for all* $\beta \in L_{\mathfrak{p}}$ *we have* $\mathrm{ord}_{\mathfrak{p}}(\beta) = e_{\mathfrak{p}}\cdot\mathrm{ord}_p(\beta)$ *). Let* q *be a rational prime such that*

$$q \nmid p\cdot(p^{f_{\mathfrak{p}}}-1) .$$

Let

$$V_j \geq \max\left(h(\alpha_j),\ f_{\mathfrak{p}}\cdot(\log p)/d \right) \quad \textit{for} \quad j = 1, \ldots, n ,$$

$$\textit{such that} \quad V_1 \leq \ldots \leq V_{n-1} , \quad V^+_{n-1} = \max(1,V_{n-1}) ,$$

$$B_0 \geq \min_{1\leq j\leq n,\, b_j\neq 0} |b_j| , \quad B_n \geq |b_n| , \quad B' \geq \max_{1\leq j\leq n-1} |b_j| ,$$

$$B \geq \max\left(|b_1|, \ldots, |b_n|, 2 \right) ,$$

$$W \geq \max\left(\log(1+\frac{3}{4\cdot n}\cdot B,\ \log B_0,\ f_{\mathfrak{p}}\cdot(\log p)/d \right) .$$

Suppose that $\mathrm{ord}_{\mathfrak{p}}(\alpha_j) = 0$ *for* $j = 1, \ldots, n$ *, that*

$$[L(\alpha_1^{1/q},\ldots,\alpha_n^{1/q}):L] = q^n , \tag{2.1}$$

that $\mathrm{ord}_p(b_n) \leq \mathrm{ord}_p(b_j)$ *for* $j = 1, \ldots, n$ *, and* $\alpha_1^{b_1}\cdot\ldots\cdot\alpha_n^{b_n} \neq 1$ *. Then*

$$\mathrm{ord}_{\mathfrak{p}}\left(\alpha_1^{b_1}\cdot\ldots\cdot\alpha_n^{b_n}-1\right) < C_1(p,n)\cdot a_1^n\cdot n^{n+5/2}\cdot q^{2\cdot n}\cdot(q-1)\cdot\log^2(n\cdot q)\cdot$$

$$(p^{f_{\mathfrak{p}}}-1)\cdot[\![2+\frac{1}{p-1}]\!]^n\cdot[\![f_{\mathfrak{p}}\cdot(\log p)/d]\!]^{-(n+2)}\cdot V_1\cdot\ldots\cdot V_n\cdot$$

$$\cdot\left(\frac{W}{6\cdot n}+\log(4\cdot d)\right)\cdot\left(\log(4\cdot d\cdot V^+_{n-1})+f_{\mathfrak{p}}\cdot(\log p)/8\cdot n\right) ,$$

where

$$a_1 = 56\cdot e/15 \quad \textit{if} \quad n \leq 7 , \quad a_1 = 8\cdot e/3 \quad \textit{if} \quad n \geq 8 ,$$

and $C_1(p,n)$ *is given by the table on the next page, with for* $p \geq 5$

$$C_1(p,n) = C'_1(p,n)\cdot[\![2+\frac{1}{p-1}]\!]^2 .$$

n	2	3	4	5	6	7	$\geqslant 8$
$C_1(2,n)$	768523	476217	373024	318871	284931	261379	2770008
$C_1(3,n)$	167881	104028	81486	69657	62243	57098	116055
$C_1'(p,n)$	87055	53944	42255	36121	32276	24584	311077

Remark. Yu [1989] gives a result in which 'independence condition' (2.1) has been removed, with more or less the same constants. This result will be easier to apply if $d \geqslant 1$.

2.5. Numerical methods.

In solving diophantine equations using computational methods from diophantine approximation theory, as we will do in Chapters 4 to 8, it is necessary to have logarithms (real, complex or p-adic) of algebraic numbers available to a large enough precision (maybe several hundreds of digits). We will not go deeply into the problems of computing such approximations, but make only a few remarks on it in this section.

To start with, the precision with which most computers (mainframes as well as personal computers) work, is insufficient for our purposes. Usually at most double precision (52 bits, equivalent to 15 decimal digits), or at best quadruple precision (112 bits, equivalent to 33 decimal digits) is standard available. This is not sufficient for our purposes, not only because we may require larger precision, but also because we want to have the rounding off errors under control, to be sure that no solution of a diophantine equation is missed by unexpected consequences of rounding off errors.

Packages for computations with arbitrary precision are available and very useful, e.g. the MP package of R.P. Brent (cf. Brent [1978]). It is not difficult, as we did, to write one's own package for simple manipulations on multi-precision numbers, such as addition, multiplication and division (cf. Knuth [1981] for efficient algorithms). To the author's knowledge, no such packages are available publicly for manipulations on p-adic numbers, but the programs are similar to those for real numbers, and thus relatively easy (though maybe laborious) to write yourself.

Computing roots of polynomials with integral coefficients can be done by

Newton's method, both in the real and the p-adic case. One should make sure that the result obtained is correct to the desired precision, not (only) by substituting the found approximation of the root into the polynomial and checking that the result is 0 within the desired precision, but (also) by theoretical error estimates for the Newton method, or by using 'interval arithmetic' (see below).

Computing logarithms can be done by the Newton method too. However, we found it easier to use the Taylor series

$$\log(1+x) = x - x^2/2 + x^3/3 - \dots ,$$

or the more rapidly converging series

$$\log\frac{1+x}{1-x} = 2\cdot\left(x + x^3/3 + x^5/5 + \dots \right) .$$

For $|x|$ very small this method works fast, whereas for larger $|x|$ the following idea works well. Compute approximations to the desired precision of $\log 1.1$, $\log 1.0001$, $\log 1.00000001$, say, and store them. Now compute $x_1 \in [1,1.1)$ and $k_1 \in \mathbb{N}_0$ such that

$$x = x_1\cdot 1.1^{k_1} ,$$

which is a matter of a few divisions of a multi-precision number with a rational number with small numerator and denominator (11 and 10) only, that can be done fast. Next, compute $x_2 \in [1,1.0001)$ and $k_2 \in \mathbb{N}_0$ such that

$$x_1 = x_2\cdot 1.0001^{k_2} ,$$

and $x_3 \in [1,1.00000001)$ and $k_3 \in \mathbb{N}_0$ such that

$$x_2 = x_3\cdot 1.00000001^{k_3} .$$

Then compute $\log x_3$ by the Taylor series, which converges very fast, and compute $\log x$ by

$$\log x = \log x_3 + k_3\cdot\log 1.00000001 + k_2\cdot\log 1.0001 + k_1\cdot\log 1.1 .$$

When computing all this, one should take care of having the rounding off errors at each addition/multiplication under control. This can e.g. be done by using 'interval arithmetic', i.e. doing all computations twice with a few more digits than actually needed, rounding off in different directions at each step. Then a sufficiently small interval is found in which the exact number lies (with mathematical certainty).

Computation of $\arctan x$ is done by the Taylor series

$$\arctan x = x - x^3/3 + x^5/5 - \ldots .$$

The number $\pi = 3.14159\ldots$ can be computed rapidly by this series for the arctan function, by the identity

$$\pi = 16\cdot\arctan 1/5 - 4\cdot\arctan 1/239 .$$

Doing p-adic arithmetic has the advantage above real arithmetic that rounding off errors do not tend to become larger, as long as one is not dividing by a number with positive p-adic order. If $\mathrm{ord}_p(x) > 0$ then $\log_p(1+x)$ can be computed by the Taylor series

$$\log_p(1+x) = x - x^2/2 + x^3/3 + \ldots ,$$

and also it may be useful to compute it by

$$\log_p \frac{1+x}{1-x} = 2\cdot\left(x + x^3/3 + x^5/5 + \ldots \right) .$$

If $x \not\equiv 0 \pmod p$ and $x \not\equiv 1 \pmod p$ then $\log_p x$ can be computed, since there exists a $k \in \mathbb{N}$ such that $x^k \equiv 1 \pmod p$, and then

$$\log_p x = \frac{1}{k}\cdot\log_p\left(1+(x^k-1)\right)$$

and the above given Taylor series can be used to compute $\log_p x$. Note that in computing the above mentioned Taylor series there will be factors p in the denominators of the terms. Hence, to find the first μ p-adic digits of $\log_p(1+x)$, it is not enough to compute only the first $\mu/\mathrm{ord}_p(x)$ terms of the Taylor series, but the first k terms must be taken into account, where k is the smallest integer satisfying

$$k\cdot\mathrm{ord}_p(x) - \log k/\log p \geqslant \mu .$$

For rapid convergence of Taylor series it is desirable to apply them only for numbers x with large p-adic order. For example,

$$\log_3 4 = 3 - 3^2/2 + 3^3/3 - \ldots$$

converges not as fast as

$$\log_3 4 = \frac{1}{3}\cdot\log_3 64 = \frac{1}{3}\cdot\left(7\cdot 3^2 - 7^2\cdot 3^4/2 + 7^3\cdot 3^6/3 - \ldots \right) ,$$

or as

$$\log_3 4 = \log_3 \frac{1+3/5}{1-3/5} = 2\cdot\left(3/5 + 3^3/3\cdot 5^3 + 3^5/5\cdot 5^5 + \ldots \right) ,$$

or as

$$\log_3 4 = \frac{1}{3}\cdot\log_3 \frac{1+7\cdot 3^2/65}{1-7\cdot 3^2/65} = \frac{2}{3}\cdot\left(7\cdot 3^2/65 + 7^3\cdot 3^6/3\cdot 65^3 \right.$$

$$\left. + 7^5\cdot 3^{10}/5\cdot 65^5 + \ldots \right) .$$

The above considerations are sufficient for efficiently performing exact computations with the L^3-algorithm, as we present it in Section 3.5. We also use the simple continued fraction algorithm in some instances. This we do as follows. Suppose we want to compute the continued fraction expansion of a real number ϑ , that we have approximated by rational numbers ϑ_1, ϑ_2 such that

$$\vartheta_1 < \vartheta < \vartheta_2 < \vartheta_1 + \varepsilon$$

for some small ε . We can compute the continued fraction expansions of ϑ_1 and ϑ_2 exactly. As far as they coincide, they coincide also with the continued fraction expansion of ϑ . If the continued fraction expansion of ϑ is needed so far that the k th convergent with denominator $q_k > X_0$ be known exactly, for a given (large) constant X_0 , then ε should be at least as small as X_0^{-2} .

Most of the computer calculations done for the research on which this book reports were performed on an IBM 3083 computer at the Centraal Rekeninstituut of the University of Leiden, using the Fortran-77 language. Whenever we give computation times, actual CPU-time on this machine is meant. Also some computations were done at a VAX 11/750 computer at the Rekencentrum of the University of Twente.

Chapter 3. Algorithms for diophantine approximation.

3.1. Introduction.

In this section we give details of the computational methods we use to reduce upper bounds for the solutions of diophantine equations. Our starting point will always be a linear form Λ that is close to 0 (in the real or p-adic sense, with the word "close" defined explicitly in terms of an inequality involving the unknowns), together with a large but explicitly known upper bound for the absolute values of the coefficients of Λ. Our aim is to reduce the upper bound by showing that there are no solutions between the new and the old upper bound.

Let $\vartheta_1, \ldots, \vartheta_n, \beta$ be given numbers, in $\mathbb{R}$, or in Ω_p, for a fixed prime p. Let $x_1, \ldots, x_n$ be unknowns in $\mathbb{Z}$. Put

$$\Lambda = \beta + \sum_{i=1}^{n} x_i \cdot \vartheta_i .$$

We classify such linear forms according to three criteria:

→ homogeneous if $\beta = 0$, inhomogeneous if $\beta \neq 0$;
→ one-dimensional if $n = 2$, multi-dimensional if $n \geqslant 3$;
→ real if $\vartheta_i \in \mathbb{R}$ for all i, p-adic if $\vartheta_i \in \Omega_p$ for all i.

The reason that the case $n = 2$ is called one-dimensional is that in the homogeneous case the linear form

$$\Lambda = x_1 \cdot \vartheta_1 + x_2 \cdot \vartheta_2$$

leads to studying the simple, one-dimensional continued fraction expansion of $-\vartheta_1/\vartheta_2$. The inhomogeneous case with $n = 1$, viz.

$$\Lambda = \beta + x \cdot \vartheta$$

is not of any interest in the real case, but it is of interest in the p-adic case. We call this the zero-dimensional case.

In the p-adic case we require that the quotients ϑ_i/ϑ_j and β/ϑ_j are in $\mathbb{Q}_p$ itself, whereas the numbers ϑ_i, β are allowed to be in some larger subfield of Ω_p.

Let c, δ be positive constants. Put $X = \max|x_i|$. Let X_0 be a (large) positive constant. In the real case we shall always assume that

$$|\Lambda| < c\cdot\exp(-\delta\cdot X)\ , \tag{3.1}$$

$$X \leqslant X_0\ . \tag{3.2}$$

Let c_1, c_2 be real constants, with $c_2 > 0$. In the p-adic case we shall assume that $x_j > 0$ for some index $j \in \{1,\ldots,n\}$, and

$$\mathrm{ord}_p(\Lambda) \geqslant c_1 + c_2\cdot x_j\ , \tag{3.3}$$

$$X \leqslant X_0\ . \tag{3.4}$$

Our aim is to find a constant X_1, of the size of $\log X_0$, such that in the real case (3.2) can be replaced by $X \leqslant X_1$, and in the p-adic case the bound $x_j \leqslant X_0$ (a consequence of (3.4)) can be improved to $x_j \leqslant X_1$.

In the forthcoming sections we will treat all cases, according to the classification given above. We insert Sections 3.4, 3.5 on the L^3-algorithm, which will be our main computational tool, Section 3.6 on finding short vectors in lattices, and Section 3.13 on certain sublattices that are useful for our applications.

3.2. Homogeneous one-dimensional approximation in the real case: continued fractions.

We first study the case

$$\Lambda = x_1\cdot\vartheta_1 + x_2\cdot\vartheta_2\ .$$

Put $\vartheta = -\vartheta_1/\vartheta_2$. We assume that ϑ is irrational. Let the continued fraction expansion of ϑ be given by

$$\vartheta = [\ a_0,\ a_1,\ a_2,\ \ldots.\ ]\ ,$$

and let the convergents p_n/q_n for $n = 0, 1, 2, \ldots$ be defined by

$$\begin{cases} p_{-1} = 1 \,, \quad p_0 = a_0 \,, \quad p_{n+1} = a_{n+1} \cdot p_n + p_{n-1} \\ q_{-1} = 0 \,, \quad q_0 = 1 \,, \quad q_{n+1} = a_{n+1} \cdot q_n + q_{n-1} \end{cases} .$$

It is well known that the convergents satisfy the inequalities

$$\frac{1}{(a_{n+1}+2) \cdot q_n^2} < \left| \vartheta - \frac{p_n}{q_n} \right| < \frac{1}{a_{n+1} \cdot q_n^2} \,, \tag{3.5}$$

and that if p/q satisfies the inequality

$$\left| \vartheta - \frac{p}{q} \right| < \frac{1}{2 \cdot q^2} \,, \tag{3.6}$$

then p/q must be one of the convergents (cf. Hardy and Wright [1979], Theorems 163, 171 and 184).

We may assume without loss of generality that $|\vartheta_1| < |\vartheta_2|$, that $x_1 > 0$, and that $(x_1, x_2) = 1$. From (3.1) it follows that there exists a number X^* such that $X \geqslant X^*$ implies $X = x_1$ and (3.6) for $(p,q) = (-x_2, x_1)$. We now have the following criteria.

<u>LEMMA 3.1.</u> (i). *If* (3.1) *and* (3.2) *hold for* x_1, x_2 *with* $X \geqslant X^*$, *then* $(-x_2, x_1) = (p_k, q_k)$ *for an index* k *that satisfies*

$$k \leqslant -1 + \log\left(\sqrt{5} \cdot X_0 + 1\right) / \log\left(\tfrac{1}{2}(1+\sqrt{5})\right) . \tag{3.7}$$

Moreover, the partial quotient a_{k+1} *satisfies*

$$a_{k+1} > -2 + |\vartheta_2| \cdot c^{-1} \cdot \exp(\delta \cdot q_k)/q_k . \tag{3.8}$$

(ii). *If for some* k *with* $q_k \geqslant X^*$

$$a_{k+1} > |\vartheta_2| \cdot c^{-1} \cdot \exp(\delta \cdot q_k)/q_k \,, \tag{3.9}$$

then (3.1) *holds for* $(-x_2, x_1) = (p_k, q_k)$.

<u>Proof.</u> (i). By $X \geqslant X^*$ and (3.6) it follows that $(-x_2, x_1) = (p_k, q_k)$ for an index k. Since q_k is at least the $(k+1)$ th Fibonacci number, (3.7) follows from $q_k = x_1 = X \leqslant X_0$. To prove (3.8), apply (3.1) and the first inequality of (3.5).
(ii). Combine (3.9) with the second inequality of (3.5). □

We may apply Lemma 3.1(i) directly, or as follows.

LEMMA 3.2. *Let*

$$A = \max(a_{k+1}) ,$$

where the maximum is taken over all indices k *satisfying* (3.7). *If* (3.1) *and* (3.2) *hold for* x_1, x_2 *with* $X \geqslant X_1$ *, then*

$$X < \frac{1}{\delta}\cdot\log\big(c\cdot(A+2)/|\vartheta_2|\big) + \frac{1}{\delta}\cdot\log X .$$

Remark. From Lemma 3.2 an upper bound for X follows. We can apply Lemma 2.1 here, but Lemma 2.1 is sharp for large b only.

Proof. (3.1) and (3.5) yield

$$(a_{n+1}+2)\cdot q_n^2 > q_n\cdot|\vartheta_2|/|\Lambda| > q_n\cdot|\vartheta_2|\cdot c^{-1}\cdot\exp(\delta\cdot X) .$$

The result follows by applying Lemma 3.1(i). □

In practice it does not often occur that A is large. Therefore this lemma is useful indeed.

Summarizing, this case comes down to computing the continued fraction of a real number to a certain precision, and establishing that it has no extremely large partial quotients. This idea has been applied in practice by Ellison [1971^b], by Cijsouw, Korlaar and Tijdeman (appendix to Stroeker and Tijdeman [1982]), and by Hunt and van der Poorten (unpublished) for solving diophantine equations, by Steiner [1977] in connection with the Syracuse ("3·N+1") problem, and by Cherubini and Walliser [1987] (using a small home computer only) for determining all imaginary quadratic number fields with class number 1. We shall use it in Chapters 4 and 5.

3.3. Inhomogeneous one-dimensional approximation in the real case: the Davenport lemma.

The next case is when Λ has the form

$$\Lambda = \beta + x_1\cdot\vartheta_1 + x_2\cdot\vartheta_2 ,$$

where $\beta \neq 0$. We then may use the so-called Davenport lemma, which was introduced by Baker and Davenport [1969]. It is, like the homogeneous case, based on the continued fraction algorithm.

Put again $\vartheta = -\vartheta_1/\vartheta_2$, and put $\psi = \beta/\vartheta_2$. Then we have

$$\frac{\Lambda}{\vartheta_2} = \psi - x_1 \cdot \vartheta + x_2 .$$

Let p/q be a convergent of ϑ with $q > X_0$. We have the following result.

<u>LEMMA 3.3. (Davenport).</u> *Suppose that, in the above notation,*

$$\|q \cdot \psi\| > 2 \cdot X_0/q , \tag{3.10}$$

(by $\|\cdot\|$ *we denote the distance to the nearest integer). Then the solutions of* (3.1), (3.2) *satisfy*

$$X < \frac{1}{\delta} \cdot \log\left(q^2 \cdot c/|\vartheta_2| \cdot X_0\right) . \tag{3.11}$$

<u>Proof.</u> From (3.5) and (3.10) we infer

$$2 \cdot X_0/q < \|q \cdot (\psi - x_1 \cdot \vartheta + x_2) + x_1 \cdot (q \cdot \vartheta - p)\| < q \cdot |\Lambda/\vartheta_2| + |x_1|/q .$$

By (3.1), (3.2), and

$$X_0 < q^2 \cdot c \cdot |\vartheta_2^{-1}| \cdot \exp(-\delta \cdot X) ,$$

this leads to (3.11). □

If (3.10) is not true for the first convergent with denominator $> X_0$, then one should try some further convergents. If q is not essentially larger than X_0 , then (3.11) yields a reduced upper bound for X of size $\log X_0$, as desired. If no q of the size of X_0 can be found that also satisfies (3.10) (a situation which is very unlikely to occur, as experiments show), then not all is lost, since then only very few exceptional possible solutions have to be checked. See Baker and Davenport [1969] for details.

Summarizing, we see that in this case the essential idea is that an extremely large solution of (3.1) and (3.2) leads to a large range of convergents p/q of ϑ for which the values of $\|q \cdot \psi\|$ are all extremely small. In practice it appears to be the case that $q \cdot \psi$ is always far enough from the nearest integer (the values of $\|q \cdot \psi\|$ seem to be distributed randomly over the interval $[0, 0.5]$). This method has been used in practice by Baker and Davenport [1969] as we already mentioned, by Ellison, Ellison, Pesek, Stahl and Stall [1972], by Steiner [1986], and by Gaál [1988]. We shall use it in Chapter 4. Note that the method that we develop in Section 3.8 for the multi-dimensional inhomogeneous case, can be used in the one-dimensional case as well, as has been demonstrated in de Weger [1989[b]].

3.4. The L^3-lattice basis reduction algorithm, theory.

To deal with linear forms with $n \geqslant 3$, a straightforward generalization of the case $n = 2$ would be to study multi-dimensional continued fractions. For a good survey of this field, see Brentjes [1981]. However, the available algorithms in this field seem not to have the desired efficiency and generality. Fortunately, since 1981 there is a useful alternative, which in a sense is also a generalization of the one-dimensional continued fraction algorithm.

In 1981, L. Lovász invented an algorithm, that has since then become known as the L^3-algorithm. It has been published in Lenstra, Lenstra and Lovász [1982], Fig. 1, p. 521. Throughout this and the next section we refer to this paper as "$\mathcal{LLL}$". The algorithm computes from an arbitrary basis of a lattice in $\mathbb{R}^n$ another basis of this lattice, a so-called *reduced* basis, which has certain nice properties (its vectors are nearly orthogonal).

The algorithm has many important applications in a variety of mathematical fields, such as the factorization of polynomials ($\mathcal{LLL}$, Lenstra [1984]), public-key cryptography (Lagarias and Odlyzko [1985]), and the disproof of the Mertens Conjecture (Odlyzko and te Riele [1985]). Of interest to us are its applications to diophantine approximation, which already had been noticed in $\mathcal{LLL}$, p. 525. The algorithm has a very good theoretical complexity (polynomial-time in the length of the input parameters), and performs also very well in practical computations.

Let $\Gamma \subset \mathbb{R}^n$ be a lattice, that is given by the basis $\underline{b}_1, \ldots, \underline{b}_n$. We introduce the concept of a *reduced* basis of Γ , according to $\mathcal{LLL}$, p. 516. The vectors $\underline{b}_i^*$ (for $i = 1, \ldots, n$) and the real numbers $\mu_{i,j}$ (for $1 \leqslant j < i \leqslant n$) are inductively defined by

$$\underline{b}_i^* = \underline{b}_i - \sum_{j=1}^{i-1} \mu_{i,j} \cdot \underline{b}_j^* , \quad \mu_{i,j} = (\underline{b}_i, \underline{b}_j^*) \,/\, (\underline{b}_j^*, \underline{b}_j^*) \; .$$

Then $\underline{b}_1^*, \ldots, \underline{b}_n^*$ is an orthogonal basis of $\mathbb{R}^n$. We call the lattice basis $\underline{b}_1, \ldots, \underline{b}_n$ of Γ *reduced* if

$$|\mu_{i,j}| \leqslant \tfrac{1}{2} \quad \text{for} \quad 1 \leqslant j < i \leqslant n \, ,$$

$$|\underline{b}_i^* + \mu_{i,i-1} \cdot \underline{b}_{i-1}^*|^2 \geqslant \tfrac{3}{4} \cdot |\underline{b}_{i-1}^*|^2 \quad \text{for} \quad 1 < i \leqslant n \; .$$

Hence a reduced basis is nearly orthogonal. For a reduced basis $\underline{b}_1, \ldots, \underline{b}_n$ we have, by $\mathcal{LLL}$ (1.7),

$$|\underline{b}_i^*| \geqslant 2^{-(n-1)/2} \cdot |\underline{b}_1| \quad \text{for} \quad i = 1, \ldots, n \; . \tag{3.12}$$

We remark that a lattice may have more than one reduced basis, and that the ordering of the basis vectors is not arbitrary. The L^3-algorithm accepts as input any basis $\underline{b}_1, \ldots, \underline{b}_n$ of Γ, and it computes a reduced basis $\underline{c}_1, \ldots, \underline{c}_n$ of that lattice. The properties of reduced bases that are of most interest to us are the following. Let $\underline{y} \in \mathbb{R}^n$ be a given point, that is not a lattice point. We denote by $\ell(\Gamma)$ the length of the shortest non-zero vector in the lattice, viz.

$$\ell(\Gamma) = \min_{\underline{0} \neq \underline{x} \in \Gamma} |\underline{x}| \; ,$$

and by $\ell(\Gamma, \underline{y})$ the distance from $\underline{y}$ to the nearest lattice point, viz.

$$\ell(\Gamma, \underline{y}) = \min_{\underline{x} \in \Gamma} |\underline{x} - \underline{y}| \; .$$

From a reduced basis lower bounds for both $\ell(\Gamma)$ and $\ell(\Gamma, \underline{y})$ can be computed, according to the following results. Lemma 3.4 is Proposition (1.11) from $\mathcal{LLL}$. We recall its proof here, to show the similarity of the proofs of Lemma's 3.4 and 3.5.

<u>LEMMA 3.4. (Lenstra, Lenstra and Lovasz [1982]).</u> *Let* $\underline{c}_1, \ldots, \underline{c}_n$ *be a reduced basis of the lattice* Γ *. Then*

$$\ell(\Gamma) \geqslant 2^{-(n-1)/2} \cdot |\underline{c}_1| \; .$$

<u>Proof.</u> Let $\underline{0} \neq \underline{x} \in \Gamma$ be the lattice point with minimal length $|\underline{x}| = \ell(\Gamma)$. Write

$$\underline{x} = \sum_{i=1}^{n} r_i \cdot \underline{c}_i = \sum_{i=1}^{n} r_i^* \cdot \underline{b}_i^* ,$$

with $r_i \in \mathbb{Z}$, $r_i^* \in \mathbb{R}$. Let i_0 be the largest index such that $r_{i_0} \neq 0$. Then, since $\underline{c}_1, \ldots, \underline{c}_i$ span the same linear space as $\underline{b}_1^*, \ldots, \underline{b}_i^*$ for all i , and $\underline{b}_{i+1}^*$ is the projection of $\underline{c}_{i+1}$ on the orthogonal complement of this linear space, it follows that $r_{i_0} = r_{i_0}^*$. Hence, by (3.12),

$$\ell(\Gamma)^2 = |\underline{x}|^2 = \sum_{i=1}^{i_0} r_i^{*2} \cdot |\underline{b}_i^*|^2 \geqslant r_{i_0}^{*2} \cdot |\underline{b}_{i_0}^*|^2 = r_{i_0}^2 \cdot |\underline{b}_{i_0}^*|^2$$

$$\geqslant |\underline{b}_{i_0}^*|^2 \geqslant 2^{-(n-1)} \cdot |\underline{c}_1|^2 . \qquad \square$$

<u>LEMMA 3.5.</u> *Let* $\underline{c}_1, \ldots, \underline{c}_n$ *be a reduced basis of the lattice* Γ *, and let* $\underline{y} = \sum_{i=1}^{n} s_i \cdot \underline{c}_i$ *for* $s_1, \ldots, s_n \in \mathbb{R}$ *, with not all* s_i *in* $\mathbb{Z}$ *. Let* i_0 *be the largest index such that* $s_{i_0} \notin \mathbb{Z}$ *. Then*

$$\ell(\Gamma, \underline{y}) \geqslant 2^{-(n-1)/2} \cdot \| s_{i_0} \| \cdot |\underline{c}_1| .$$

<u>Proof.</u> Let $\underline{x} \in \Gamma$ be the lattice point nearest to $\underline{y}$. So $|\underline{x}-\underline{y}| = \ell(\Gamma, \underline{y})$. Write

$$\underline{x} = \sum_{i=1}^{n} r_i \cdot \underline{c}_i = \sum_{i=1}^{n} r_i^* \cdot \underline{b}_i^* , \quad \underline{y} = \sum_{i=1}^{n} s_i \cdot \underline{c}_i = \sum_{i=1}^{n} s_i^* \cdot \underline{b}_i^* ,$$

with $r_i \in \mathbb{Z}$, $r_i^*, s_i, s_i^* \in \mathbb{R}$. Let i_1 be the largest index such that $r_{i_1} \neq s_{i_1}$. Then, reasoning as in the proof of Lemma 3.4, we find

$$r_{i_1} - s_{i_1} = r_{i_1}^* - s_{i_1}^* .$$

Using (3.12) it follows that

$$\ell(\Gamma, \underline{y})^2 \geqslant (r_{i_1} - s_{i_1})^2 \cdot |\underline{b}_{i_1}^*|^2 \geqslant (r_{i_1} - s_{i_1})^2 \cdot 2^{-(n-1)} \cdot |\underline{c}_1|^2 .$$

Obviously, $i_1 \geqslant i_0$. If $i_1 = i_0$ the result follows at once. If $i_1 > i_0$ then $s_{i_1} \in \mathbb{Z}$, $s_{i_1} \neq r_{i_1}$, hence $|r_{i_1} - s_{i_1}| \geqslant 1$, and the result follows. $\square$

The above lemma is rather weak in the extraordinary situation that s_{i_0} is extremely close to an integer. If one of the other s_i is not close to an integer, we can apply the following variant.

<u>LEMMA 3.6.</u> *Let* $\underline{c}_1, \ldots, \underline{c}_n$ *be a reduced basis of the lattice* Γ *, and let* $\underline{y} = \sum_{i=1}^{n} s_i \cdot \underline{c}_i$ *for* $s_1, \ldots, s_n \in \mathbb{R}$ *, with not all* s_i *in* $\mathbb{Z}$ *. Suppose that there is an index* i_0 *and constants* δ_1 *,* $0 < \delta_2 \leqslant \frac{1}{2}$ *such that*

$$\|s_i\| \leqslant \delta_1 \quad \text{for} \quad i = i_0+1, \ldots, n ,$$

$$\|s_{i_0}\| \geqslant \delta_2 .$$

Then

$$\ell(\Gamma, \underline{y}) \geqslant 2^{-(n-1)/2} \cdot \delta_2 \cdot |\underline{c}_1| - (n-i_0) \cdot \delta_1 \cdot \max_{i>i_0} |\underline{c}_i| .$$

<u>Proof.</u> With notation as in the proof of Lemma 3.5, let t_i be the integer nearest to s_i , for $i \geqslant i_0 + 1$, and $t_i = s_i$ for $i \leqslant i_0$. Put

$$\underline{z} = \sum_{i=1}^{n} t_i \cdot \underline{c}_i = \sum_{i=1}^{n} t_i^* \cdot \underline{b}_i^*$$

with $t_i^* \in \mathbb{R}$. Let i_1 be the largest index such that $r_{i_1} \neq t_{i_1}$. Then

$$r_{i_1} - t_{i_1} = r_{i_1}^* - t_{i_1}^* .$$

We have

$$\ell(\Gamma, \underline{y}) = |\underline{x}-\underline{y}| \geqslant |\underline{x}-\underline{z}| - |\underline{z}-\underline{y}| .$$

Now,

$$|\underline{z}-\underline{y}| \leqslant \sum_{i=i_0+1}^{n} |s_i - t_i| \cdot |\underline{c}_i| \leqslant (n-i_0) \cdot \delta_1 \cdot \max_{i>i_0} |\underline{c}_i| ,$$

and, using (3.12),

$$|\underline{x}-\underline{z}|^2 = \sum_{i=1}^{n} (r_i^* - t_i^*)^2 \cdot |\underline{b}_i^*|^2 \geqslant (r_{i_1}^* - t_{i_1}^*)^2 \cdot |\underline{b}_{i_1}^*|^2$$

$$\geqslant (r_{i_1} - t_{i_1})^2 \cdot 2^{-(n-1)} \cdot |\underline{c}_1|^2 .$$

Obviously, $i_1 \geqslant i_0$. If $i_1 = i_0$ the result follows. If $i_1 > i_0$ then

$t_{i_1} \in \mathbb{Z}$, $t_{i_1} \neq r_{i_1}$, hence $|r_{i_1} - t_{i_1}| \geqslant 1 > \delta_2$, and the result follows. $\square$

Remark. Babai [1986] showed that the L^3-algorithm can be used to find a lattice point $\underline{x}$ with $|\underline{x}-\underline{y}| \leqslant c \cdot \ell(\Gamma, \underline{y})$ for a constant c depending on the dimension of the lattice only. This result can also be used instead of Lemma 3.5 or 3.6.

3.5. The L^3-lattice basis reduction algorithm, practice.

Below (in Fig. 1) we describe the variant of the L^3-algorithm that we use in this monograph to solve diophantine equations. This variant has been designed to work with integers only, so that rounding-off errors are avoided completely. In the algorithm as stated in $\mathscr{LLL}$, Fig. 1, p. 521, non-integral rational numbers may occur, even if the input parameters are all integers.

Let $\Gamma \subset \mathbb{Z}^n$ be a lattice with basis vectors $\underline{b}_1, \ldots, \underline{b}_n$. Define $\underline{b}_i^*$, μ_{ij}, d_i as in $\mathscr{LLL}$ (1.2), (1.3), (1.24), respectively. The d_i can be used as denominators for all numbers that appear in the original algorithm ($\mathscr{LLL}$, p. 523). Thus, put for all relevant indices i, j

$$\begin{aligned} \underline{c}_i &= d_{i-1} \cdot \underline{b}_i^* , \\ \lambda_{i,j} &= d_j \cdot \mu_{i,j} . \end{aligned} \tag{3.13}$$

They are integral, by $\mathscr{LLL}$ (1.28), (1.29). Notice that, with $B_i = |\underline{b}_i^*|^2$,

$$d_i = d_{i-1} \cdot B_i . \tag{3.14}$$

We can now rewrite the algorithm in terms of $\underline{c}_i$, d_i, $\lambda_{i,j}$ in stead of $\underline{b}_i^*$, B_i, $\mu_{i,j}$, thus eliminating all non-integral rationals. We give this variant of the L^3-algorithm in Fig. 1. All the lines in this variant are evident from applying (3.13) and (3.14) to the corresponding lines in the original algorithm, except the lines (A), (B) and (C), which will be explained below.

We added a few lines to the algorithm, in order to compute the matrix of the transformation from the initial to the reduced basis. Let $\mathscr{B}$ be the matrix with column vectors $\underline{b}_1, \ldots, \underline{b}_n$, the initial basis of the lattice Γ, which is the input for the algorithm. We say: $\mathscr{B}$ is the matrix *associated to* the basis $\underline{b}_1, \ldots, \underline{b}_n$. Let $\mathscr{C}$ be the matrix associated to the reduced

$$
\begin{array}{ll}
 & d_0 := 1 \; ; \\
(A) & \left.\begin{array}{l} \underline{c}_i := \underline{b}_i \; ; \\ \left.\begin{array}{l} \lambda_{i,j} := (\underline{b}_i, \underline{c}_j) \; ; \\ \underline{c}_i := (d_j \cdot \underline{c}_i - \lambda_{i,j} \cdot \underline{c}_j)/d_{j-1} \end{array}\right\} \text{ for } j=1,\ldots,i-1 \; ; \\ d_i := (\underline{c}_i, \underline{c}_i)/d_{i-1} \end{array}\right\} \text{ for } i=1,\ldots,n \; ; \\
 & k := 2 \; ; \\
(1) & \text{perform } (*) \text{ for } \ell = k-1 \; ; \\
 & \text{if } 4 \cdot d_{k-2} \cdot d_k < 3 \cdot d_{k-1}^2 - 4 \cdot \lambda_{k,k-1}^2 \text{ go to } (2) \; ; \\
 & \text{perform } (*) \text{ for } \ell = k-2, \ldots, 1 \; ; \\
 & \text{if } k = n \text{ terminate } ; \\
 & k := k+1 \; ; \text{ go to } (1) \; ; \\
(2) & \begin{pmatrix} \underline{b}_{k-1} \\ \underline{b}_k \end{pmatrix} := \begin{pmatrix} \underline{b}_k \\ \underline{b}_{k-1} \end{pmatrix} \; ; \\
 & \begin{pmatrix} \underline{u}_{k-1} \\ \underline{u}_k \end{pmatrix} := \begin{pmatrix} \underline{u}_k \\ \underline{u}_{k-1} \end{pmatrix} \; ; \quad \begin{pmatrix} \underline{v}'^T_{k-1} \\ \underline{v}'^T_k \end{pmatrix} := \begin{pmatrix} \underline{v}'^T_k \\ \underline{v}'^T_{k-1} \end{pmatrix} \; ; \\
 & \begin{pmatrix} \lambda_{k-1,j} \\ \lambda_{k,j} \end{pmatrix} := \begin{pmatrix} \lambda_{k,j} \\ \lambda_{k-1,j} \end{pmatrix} \text{ for } j = 1, \ldots, k-2 \; ; \\
(B) & \begin{pmatrix} \lambda_{i,k-1} \\ \lambda_{i,k} \end{pmatrix} := \left(\lambda_{i,k-1} \cdot \begin{pmatrix} \lambda_{k,k-1} \\ d_k \end{pmatrix} + \lambda_{i,k} \cdot \begin{pmatrix} d_{k-2} \\ -\lambda_{k,k-1} \end{pmatrix} \right) / \, d_{k-1} \\
 & \hfill \text{for } i = k+1, \ldots, n \; ; \\
(C) & d_{k-1} := \left(d_{k-2} \cdot d_k + \lambda_{k,k-1}^2 \right) / \, d_{k-1} \; ; \\
 & \text{if } k > 2 \text{ then } k := k-1 \; ; \\
 & \text{go to } (1) \; ; \\
(*) & \text{if } 2 \cdot |\lambda_{k,\ell}| > d_\ell \text{ then} \\
 & \left\{\begin{array}{l} r := \text{integer nearest to } \lambda_{k,\ell}/d_\ell \; ; \\ \underline{b}_k := \underline{b}_k - r \cdot \underline{b}_\ell \; ; \quad \underline{u}_k := \underline{u}_k - r \cdot \underline{u}_\ell \; ; \quad \underline{v}'^T_\ell := \underline{v}'^T_\ell + r \cdot \underline{v}'^T_k \; ; \\ \lambda_{k,j} := \lambda_{k,j} - r \cdot \lambda_{\ell,j} \text{ for } j = 1, \ldots, \ell-1 \; ; \\ \lambda_{k,\ell} := \lambda_{k,\ell} - r \cdot d_\ell \; . \end{array}\right.
\end{array}
$$

Figure 1. Variant of the L^3-algorithm.

basis $\underline{c}_1, \ldots, \underline{c}_n$, which the algorithm delivers as output. Then we define this transformation matrix $\mathcal{V}$ by

$$\mathcal{C} = \mathcal{B}\cdot\mathcal{V} .$$

More generally, let $\mathcal{U}$ be the matrix of a transformation from some $\mathcal{B}_0$ to $\mathcal{B}$, so $\mathcal{B} = \mathcal{B}_0\cdot\mathcal{U}$. Denote the column vectors of $\mathcal{U}$ by $\underline{u}_1, \ldots, \underline{u}_n$, and the row vectors of $\mathcal{U}^{-1}$ by $\underline{v}'^T_1, \ldots, \underline{v}'^T_n$. We feed the algorithm with $\mathcal{U}$ and $\mathcal{U}^{-1}$ as well. All manipulations in the algorithm done on the $\underline{b}_i$ are also done on the $\underline{u}_i$, and the $\underline{v}'^T_i$ are adjusted accordingly. This does not affect the computation time seriously. The algorithm now gives as output matrices $\mathcal{C}$, $\mathcal{U}'$ and $\mathcal{U}'^{-1}$, such that $\mathcal{C}$ is associated to a reduced basis, $\mathcal{C} = \mathcal{B}\cdot\mathcal{V}$, and $\mathcal{U}' = \mathcal{U}\cdot\mathcal{V}$. Note that $\mathcal{V}$ is not computed explicitly, unless $\mathcal{U} = \mathcal{I}$ (the unit matrix), in which case $\mathcal{U}' = \mathcal{V}$. It follows that

$$\mathcal{C} = \mathcal{B}\cdot\mathcal{U}^{-1}\cdot\mathcal{U}' = \mathcal{B}_0\cdot\mathcal{U}' ,$$

so $\mathcal{U}'$ is the matrix of the transformation from $\mathcal{B}_0$ to $\mathcal{C}$. Note that if $\mathcal{B}_0^{-1}$ is known, then it is not much extra effort to compute $\mathcal{C}^{-1}$ as well.

We now explain why lines (A), (B) and (C) are correct.

(A): From $\mathcal{LLL}$ (1.2) it follows that

$$\underline{c}_i = d_{i-1}\cdot\underline{b}_i - \sum_{k=1}^{i-1} \frac{d_{i-1}}{d_{k-1}\cdot d_k}\cdot\lambda_{i,k}\cdot\underline{c}_k .$$

Define for $j = 0, 1, \ldots, i-1$

$$\underline{c}_i(j) = d_j\cdot\underline{b}_i - \sum_{k=1}^{j} \frac{d_j}{d_{k-1}\cdot d_k}\cdot\lambda_{i,k}\cdot\underline{c}_k .$$

Then $\underline{c}_i(0) = \underline{b}_i$, and $\underline{c}_i(i-1) = \underline{c}_i$. The $\underline{c}_i(j)$ is exactly the vector computed in (A) at the j th step, since

$$\frac{d_j\cdot\underline{c}_i(j-1) - \lambda_{i,j}\cdot\underline{c}_j}{d_{j-1}}$$

$$= d_j\cdot\underline{b}_i - \sum_{k=1}^{j-1} \frac{d_j}{d_{k-1}\cdot d_k}\cdot\lambda_{i,k}\cdot\underline{c}_k - \frac{d_j}{d_{j-1}\cdot d_j}\cdot\lambda_{i,j}\cdot\underline{c}_j = \underline{c}_i(j) .$$

This explains the recursive formula in line (A). It remains to show that the occurring vectors $\underline{c}_i(j)$ are integral. This follows from

$$d_j \cdot \sum_{k=1}^{j} \frac{1}{d_{k-1} \cdot d_k} \cdot \lambda_{i,k} \cdot \underline{c}_k = d_j \cdot \sum_{k=1}^{j} \mu_{i,k} \cdot \underline{b}_k^* ,$$

which is integral by $\mathcal{LLL}$ p. 523, ℓ. 11.

(B), (C): Notice that the third and fourth line, starting from label (2), in the original algorithm, are independent of the first, second and fifth line. Thus a permutation of these lines is allowed. We rewrite the first, second and fifth line as follows (where we indicate variables that have been changed with a prime sign):

$$B'_{k-1} := B_k + \mu^2_{k,k-1} \cdot B_{k-1} ; \tag{3.15}$$

$$B'_k := B_{k-1} \cdot B_k / B'_{k-1} ; \tag{3.16}$$

$$\mu'_{k,k-1} := \mu_{k,k-1} \cdot B_{k-1} / B'_{k-1} ; \tag{3.17}$$

$$\mu'_{i,k-1} := \mu'_{k,k-1} \cdot \mu_{i,k-1} + (1-\mu_{k,k-1} \cdot \mu'_{k,k-1}) \cdot \mu_{i,k} ; \tag{3.18}$$

$$\mu'_{i,k} := \mu_{i,k-1} - \mu_{k,k-1} \cdot \mu_{i,k} ; \tag{3.19}$$

where (3.18) and (3.19) hold for $i = k+1, \ldots, n$. The d_i remain unchanged for $i = 0, 1, \ldots, k-2$, and by (3.16) also for $i = k$. Now, (3.15) is equivalent to

$$\frac{d'_{k-1}}{d_{k-2}} = \frac{d_k}{d_{k-1}} + \frac{\lambda^2_{k,k-1}}{d^2_{k-1}} \cdot \frac{d_{k-1}}{d_{k-2}} , \tag{3.20}$$

which explains (C). From (3.17) we find

$$\frac{\lambda'_{k,k-1}}{d'_{k-1}} = \frac{\lambda_{k,k-1}}{d_{k-1}} \cdot \frac{d_{k-1}}{d_{k-2}} \cdot \frac{d'_{k-2}}{d'_{k-1}} ,$$

hence $\lambda_{k,k-1}$ remains unchanged. From (3.18) we obtain

$$\frac{\lambda'_{i,k-1}}{d'_{k-1}} = \frac{\lambda_{k,k-1}}{d'_{k-1}} \cdot \frac{\lambda_{i,k-1}}{d_{k-1}} + \left(1 - \frac{\lambda_{k,k-1}}{d_{k-1}} \cdot \frac{\lambda_{k,k-1}}{d'_{k-1}} \right) \cdot \frac{\lambda_{i,k}}{d_k} ,$$

whence, by multiplying by $d_{k-1} \cdot d'_{k-1}$ and using (3.20),

$$\begin{aligned} d_{k-1} \cdot \lambda'_{i,k-1} &= \lambda_{k,k-1} \cdot \lambda_{i,k-1} + (d_{k-1} \cdot d'_{k-1} - \lambda^2_{k,k-1}) \cdot \frac{\lambda_{i,k}}{d_k} \\ &= \lambda_{k,k-1} \cdot \lambda_{i,k-1} + d_{k-2} \cdot \lambda_{i,k} . \end{aligned}$$

Finally, from (3.19) we see

$$\frac{\lambda'_{i,k}}{d_k} = \frac{\lambda_{i,k-1}}{d_{k-1}} - \frac{\lambda_{k,k-1}}{d_{k-1}} \cdot \frac{\lambda_{i,k}}{d_k} ,$$

and (B) follows.

In our applications we often have a lattice Γ , of which a basis is given such that the associated matrix, $\mathcal{A}$ say, has the special form

$$\mathcal{A} = \begin{pmatrix} 1 & & & \emptyset \\ & \ddots & & \\ \emptyset & & 1 & \\ \Theta_1 & \dots & \Theta_{n-1} & \Theta_n \end{pmatrix} ,$$

where the Θ_i are large integers, that may have several hundreds of decimal digits. We can compute a reduced basis of this lattice directly, using the matrix $\mathcal{A}$ itself as input for the L^3-algorithm. But it may save time and space to split up the computation into several steps with increasing accuracy, as follows.

Let k be a natural number (the number of steps), and let ℓ be a natural number such that the Θ_i have about $k \cdot \ell$ (decimal) digits. For $i = 1, \dots, n$ and $j = 1, \dots, k$ put

$$\Theta_i^{(j)} = [\Theta_i / 10^{\ell \cdot (k-j)}] ,$$

and define $\Psi_i^{(j)}$ by

$$\Theta_i^{(j+1)} = 10^{\ell} \cdot \Theta_i^{(j)} + \Psi_i^{(j)} .$$

Thus, the $\Psi_i^{(j)}$ are blocks of ℓ consecutive digits of Θ_i . Define for the relevant j the $n \times n$ matrices

$$\mathcal{A}_j = \begin{pmatrix} 1 & & & \emptyset \\ & \ddots & & \\ \emptyset & & 1 & \\ \Theta_1^{(j)} & \dots & \Theta_{n-1}^{(j)} & \Theta_n^{(j)} \end{pmatrix} , \quad \mathcal{D}_j = \begin{pmatrix} & & \\ & \emptyset & \\ & & \\ \Psi_1^{(j)} & \dots & \Psi_n^{(j)} \end{pmatrix} ,$$

$$\mathcal{E} = \begin{pmatrix} 1 & & & \emptyset \\ & \ddots & & \\ \emptyset & & 1 & \\ & & & 10^{\ell} \end{pmatrix} .$$

Then it follows at once that

$$\mathcal{A}_{j+1} = \mathcal{E}\cdot\mathcal{A}_j + \mathcal{D}_j \ .$$

Notice that $\mathcal{A}_k = \mathcal{A}$, since $\Theta_i^{(k)} = \Theta_i$. Put $\mathcal{U}_0 = \mathcal{I}$, $\mathcal{B}_1 = \mathcal{A}_1$. For some $j \geqslant 1$ let $\mathcal{B}_j$ and $\mathcal{U}_{j-1}$ be known matrices. Then we apply the L^3-algorithm to $\mathcal{B} = \mathcal{B}_j$, $\mathcal{U} = \mathcal{U}_{j-1}$, and $\mathcal{U}^{-1}$. We thus find matrices $\mathcal{C}_j$, $\mathcal{U}_j$, and $\mathcal{U}_j^{-1}$ such that

$$\mathcal{C}_j = \mathcal{B}_j\cdot\mathcal{U}_{j-1}^{-1}\cdot\mathcal{U}_j \ .$$

Now put

$$\mathcal{B}_{j+1} = \mathcal{E}\cdot\mathcal{C}_j + \mathcal{D}_j\cdot\mathcal{U}_j \ .$$

By induction $\mathcal{B}_j$, $\mathcal{C}_j$ and $\mathcal{U}_j$ are defined for $j = 1, \ldots, k$. Note that

$$\mathcal{B}_{j+1}\cdot\mathcal{U}_j^{-1} = \mathcal{E}\cdot\mathcal{B}_j\cdot\mathcal{U}_{j-1}^{-1} + \mathcal{D}_j \ ,$$

so the $\mathcal{B}_j\cdot\mathcal{U}_{j-1}^{-1}$ satisfy the same recursive relation as the $\mathcal{A}_j$. Since $\mathcal{B}_1\cdot\mathcal{U}_0^{-1} = \mathcal{A}_1$, we have $\mathcal{B}_j\cdot\mathcal{U}_{j-1}^{-1} = \mathcal{A}_j$ for all j . Hence

$$\mathcal{C}_j = \mathcal{B}_j\cdot\mathcal{U}_{j-1}^{-1}\cdot\mathcal{U}_j = \mathcal{A}_j\cdot\mathcal{U}_j \ ,$$

and it follows that $\mathcal{C}_k$ and $\mathcal{A}_k$ are associated to bases of the same lattice, which is Γ . Moreover, since $\mathcal{C}_k$ is output of the L^3-algorithm, it is associated to a reduced basis of Γ .

Let us now analyse the computation time. For a matrix $\mathcal{M}$ we denote by $L(\mathcal{M})$ the maximal number of (decimal) digits of its entries. If the L^3-algorithm is applied to a matrix $\mathcal{B}$, with as output a matrix $\mathcal{C}$, then according to the experiences of Lenstra, Odlyzko (cf. Lenstra [1984], p. 7) and ourselves, the computation time is proportional to $L(\mathcal{B})^3$ in practice. Since $\mathcal{C}$ is associated to a reduced basis, we assume that

$$L(\mathcal{C}) \cong {}^{10}\log(\det \Gamma)/n \ .$$

In our situation, $L(\mathcal{A}_j) \cong \ell\cdot j$, $L(\mathcal{D}_j) \cong \ell$, and by $\det \mathcal{C}_j = \det \mathcal{A}_j = \Theta_n^{(j)}$ we have $L(\mathcal{C}_j) \cong \ell\cdot j/n$. Put $\mathcal{C}_j = (c_{i,h}^{(j)})$, $\mathcal{U}_j = (u_{i,h}^{(j)})$. Then by $\mathcal{C}_j = \mathcal{A}_j\cdot\mathcal{U}_j$ and the special shape of $\mathcal{A}_j$ we have $c_{i,h}^{(j)} = u_{i,h}^{(j)}$ for $i = 1, \ldots, n-1$ and $h = 1, \ldots, n$, and

$$u_{n,h}^{(j)} = \Big(- c_{1,h}^{(j)}\cdot\Theta_1^{(j)} - \ldots - c_{n-1,h}^{(j)}\cdot\Theta_{n-1}^{(j)} + c_{n,h}^{(j)} \Big)/\Theta_n^{(j)} \ .$$

It follows that $L(\mathcal{U}_j) \cong L(\mathcal{C}_j)$. So

$$L(\mathcal{B}_j) \cong \max\left(L(\mathcal{E}\cdot\mathcal{C}_{j-1}),\ L(\mathcal{D}_{j-1}\cdot\mathcal{U}_{j-1}) \right) \cong \ell + \ell\cdot(j-1)/n \ .$$

Instead of applying the L^3-algorithm once with $\mathcal{A}$ as input, we apply it k times, with $\mathcal{B}_1, \ldots, \mathcal{B}_k$ as input. Thus we reduce the computation time by a factor

$$\frac{L(\mathcal{A})^3}{\sum_{j=1}^{k} L(\mathcal{B}_j)^3} \cong \frac{(\ell\cdot k)^3}{\sum_{j=1}^{k} \ell^3\cdot\left(1+\frac{j-1}{n}\right)^3} = \frac{k^3\cdot n^3}{\sum_{j=0}^{k-1} (n+j)^3} .$$

For k between $2.5\cdot n$ and $3\cdot n$ this expression is maximal, about $0.4\cdot n^2$. So the reduction in computation time is considerable (a factor 10 already for $n = 5$). The storage space that is required is also reduced, since the largest numbers that appear in the input have $\ell\cdot\left(1+(k-1)/n\right)$ instead of $\ell\cdot k$ digits.

3.6. Finding all short lattice points: the Fincke and Pohst algorithm.

Sometimes it is not sufficient to have only a lower bound for $\ell(\Gamma)$ or $\ell(\Gamma,\underline{y})$. It may be useful to know exactly all vectors $\underline{x} \in \Gamma$ such that $|\underline{x}| \leqslant C$ or $|\underline{x}-\underline{y}| \leqslant C$ for a given constant C . There exists an efficient algorithm for finding all the solutions to these problems. This algorithm was devised by Fincke and Pohst [1985], cf. their (2.8) and (2.12). We give a description of this algorithm below.

The input of the algorithm is a matrix $\mathcal{B}$ whose column vectors span the lattice Γ , and a constant $C > 0$. The output is a list of all lattice points $\underline{x} \in \Gamma$ with $|\underline{x}| \leqslant C$, apart from $\underline{x} = \underline{0}$. We give the algorithm in Figure 2. We use the notation $\mathcal{X} = (x_{ij})$ for matrices $\mathcal{X} = \mathcal{A}, \mathcal{B}, \mathcal{R}, \mathcal{S}, \mathcal{U}$, and $\underline{x}_i$ for the column vectors of $\mathcal{X}$.

The algorithm can also be used for finding all vectors $\underline{x} \in \Gamma$ of which the distance to a given non-lattice point $\underline{y}$ is at most a given constant C . Namely, let

$$\underline{y} = \sum_{i=1}^{n} s_i\cdot\underline{b}_i \ ,$$

and let r_i be the integer nearest to s_i for all i . Put

$\mathcal{A} := \mathcal{B}^T \cdot \mathcal{B}$;

$q_{ij} := a_{ij}$ for $1 \leqslant i \leqslant j \leqslant n$;

$q_{ji} := q_{ij}$, $q_{ij} := q_{ij}/q_{ii}$ for $1 \leqslant i < j \leqslant n$;

$q_{k\ell} := q_{k\ell} - q_{ki} \cdot q_{i\ell}$ for $i+1 \leqslant k \leqslant \ell \leqslant n$ for $1 \leqslant i \leqslant n$;

$r_{ii} := \sqrt{q_{ii}}$ for $1 \leqslant i \leqslant n$;

$r_{ij} := r_{ii} \cdot q_{ij}$, $r_{ji} := 0$ for $1 \leqslant j < i \leqslant n$;

compute $\mathcal{R}^{-1}$;

compute a row-reduced version $\mathcal{S}^{-1}$ of $\mathcal{R}^{-1}$, and $\mathcal{U}$, $\mathcal{U}^{-1}$ such that $\mathcal{S}^{-1} = \mathcal{U}^{-1} \cdot \mathcal{R}^{-1}$;

compute $\mathcal{S} = \mathcal{R} \cdot \mathcal{U}4$;

determine a permutation π such that $|\underline{s}_{\pi(1)}| \geqslant \ldots \geqslant |\underline{s}_{\pi(n)}|$, let $\mathcal{S}'$ be the matrix with columns $\underline{s}_{\pi^{-1}(i)}$ for $i = 1,\ldots,n$;

$\mathcal{A} := \mathcal{S}'^T \cdot \mathcal{S}'$;

$q_{ij} := a_{ij}$ for $1 \leqslant i \leqslant j \leqslant n$;

$q_{ji} := q_{ij}$, $q_{ij} := q_{ij}/q_{ii}$ for $1 \leqslant i < j \leqslant n$;

$q_{k\ell} := q_{k\ell} - q_{ki} \cdot q_{i\ell}$ for $i+1 \leqslant k \leqslant \ell \leqslant n$ for $1 \leqslant i \leqslant n$;

$i := n$;

$T_i := C$;

$U_i := 0$;

(1) $Z := \sqrt{(T_i / q_{ii})}$;

$UB(x_i) := \lfloor Z - U_i \rfloor$;

$x_i := \lceil -Z - U_i \rceil - 1$;

(2) $x_i := x_i + 1$;

if $x_i \leqslant UB(x_i)$, go to (4) ;

(3) $i := i + 1$;

go to (2) ;

(4) if $i = 1$, go to (5) ;

$i := i - 1$;

$U_i := \sum_{j=i+1}^{m} q_{ij} \cdot x_j$;

$T_i := T_{i+1} - q_{i+1,i+1} \cdot (x_{i+1} + U_{i+1})^2$;

go to (1) ;

(5) if $x_i = 0$ for $1 \leqslant i \leqslant n$, terminate ;

compute and print $\underline{x} = \mathcal{U} \cdot (x_{\pi^{-1}(1)}, \ldots, x_{\pi^{-1}(n)})^T$;

go to (2).

Figure 2. The Fincke and Pohst Algorithm.

$$\underline{z} = \sum_{i=1}^{n} r_i \cdot \underline{b}_i \ .$$

Then $|\underline{y}-\underline{z}| < C'$ for some constant C' ($C' = \frac{n}{2}\cdot\sum|\underline{b}_i|$ will do). Since $\underline{z} \in \Gamma$ it suffices to search for all lattice points $\underline{u}$ with $|\underline{u}| \leqslant C + C'$, and compute for each such $\underline{u}$ also $\underline{x} = \underline{z} + \underline{u}$, since $|\underline{x}-\underline{y}| < C$ implies

$$|\underline{u}| \leqslant |\underline{x}-\underline{y}| + |\underline{y}-\underline{z}| \leqslant C + C' \ .$$

3.7. Homogeneous multi-dimensional approximation in the real case: real approximation lattices.

Let the linear form Λ have the form

$$\Lambda = \sum_{i=1}^{n} x_i \cdot \vartheta_i \ .$$

We assume that $n \geqslant 2$. The case $n = 2$ has already been discussed in Section 3.2, but the method of this section works also for $n = 2$. In fact, it is in this case essentially the same method.

Let C be a large enough integer, that is of the order of magnitude of X_0^n . Let $\gamma \in \mathbb{N}$ be a constant (we will explain its use later). We define the *approximation lattice* Γ by the matrix

$$\mathcal{B} = \begin{pmatrix} \gamma & & & & \emptyset \\ & \cdot & & & \\ & & \cdot & & \\ \emptyset & & & \cdot & \\ & & & \gamma & \\ [\gamma\cdot C\cdot\vartheta_1] & \dots & & [\gamma\cdot C\cdot\vartheta_{n-1}] & [\gamma\cdot C\cdot\vartheta_n] \end{pmatrix} ,$$

of which the column vectors $\underline{b}_1, \dots, \underline{b}_n$ are a basis of the lattice. Then Γ is a sublattice of $\mathbb{Z}^n$ of determinant $\gamma^{n-1}\cdot[\gamma\cdot C\cdot\vartheta_n]$, which is of size C . A lattice point $\underline{x}$ has the form

$$\underline{x} = \sum_{i=1}^{n} x_i \cdot \underline{b}_i = (\gamma\cdot x_1, \dots, \gamma\cdot x_{n-1}, \tilde{\Lambda})^T ,$$

where the x_i are integers, and

$$\tilde{\Lambda} = \sum_{i=1}^{n} x_i \cdot [\gamma\cdot C\cdot\vartheta_i] \ .$$

Clearly, $\tilde{\Lambda}$ is close to $\gamma\cdot C\cdot\Lambda$. The length of the vector $\underline{x}$ now measures both X_0 and $|\Lambda|$, which are exactly the two numbers we want to balance with each other. Heuristics (cf. Section 1.3) tell us that in a generic case we expect $|\Lambda| \cong X_0^{-n}$. We now can prove easily the following useful lemma.

<u>LEMMA 3.7.</u> *Let* X_1 *be a positive number such that*

$$\ell(\Gamma) \geqslant \sqrt{\left((n+1)^2+(n-1)\cdot\gamma^2\right)}\cdot X_1 \ . \tag{3.21}$$

Then (3.1) *has no solutions with*

$$\frac{1}{\delta}\cdot\log(\gamma\cdot C\cdot c/X_1) \leqslant X \leqslant X_1 \ . \tag{3.22}$$

<u>Remark.</u> We apply this lemma for $X_1 = X_0$. If condition (3.21) then fails, we must take a larger constant C . If it holds for a constant C of the size X_0^n , then (3.22) yields a reduced lower bound for X of size $\log X_0$.

<u>Proof.</u> Let $x_1, \ldots, x_n$ be a solution of (3.1) with $0 < X \leqslant X_1$. Consider the lattice point

$$\underline{x} = \sum_{i=1}^{n} x_i\cdot\underline{b}_i = \left(\gamma\cdot x_1, \ \ldots, \ \gamma\cdot x_{n-1}, \ \tilde{\Lambda}\right)^T ,$$

with $\tilde{\Lambda}$ as above. Then

$$|\underline{x}|^2 = \gamma^2\cdot\sum_{i=1}^{n-1} x_i^2 + \tilde{\Lambda}^2 \leqslant (n-1)\cdot\gamma^2\cdot X_1^2 + \tilde{\Lambda}^2 ,$$

and

$$|\tilde{\Lambda}-\gamma\cdot C\cdot\Lambda| \leqslant \sum_{i=1}^{n} |x_i|\cdot|[\gamma\cdot C\cdot\vartheta_i]-\gamma\cdot C\cdot\vartheta_i| \leqslant \sum_{i=1}^{n} |x_i| , \tag{3.23}$$

which is $\leqslant n\cdot X_1$. By (3.1), (3.21) and the definition of $\ell(\Gamma)$ we have

$$\gamma\cdot C\cdot c\cdot\exp(-\delta\cdot X) > |\gamma\cdot C\cdot\Lambda| \geqslant |\tilde{\Lambda}| - |\tilde{\Lambda}-\gamma\cdot C\cdot\Lambda|$$

$$\geqslant \sqrt{\left(\ell(\Gamma)^2-(n-1)\cdot\gamma^2\cdot X_1^2\right)} - n\cdot X_1 \geqslant X_1 ,$$

and (3.22) follows at once. □

Condition (3.21) can be checked by computing a reduced basis of the lattice Γ by the L^3-algorithm, and applying Lemma 3.4. The parameter γ is used to keep the "rounding-off error"

$$|[\gamma\cdot C\cdot\vartheta_i]-\gamma\cdot C\cdot\vartheta_i|$$

relatively small. This is of importance only if C is not very large, usually only if one wants to make a further reduction step after the first step has already been made. For large C, simply take $\gamma = 1$.

It may be necessary, if C is not very large, to use a more refined method of reducing the upper bound. To do so, we use the following lemma, which is a slight refinement of Lemma 3.7, together with the algorithm of Fincke and Pohst (cf. Section 3.6). It is particularly useful in the situation that one has different upper bounds for the $|x_i|$ for different i.

LEMMA 3.8. *Suppose that for a solution of* (3.1)

$$|\tilde{\Lambda}| > \sum_{i=1}^{n} |x_i| \tag{3.24}$$

holds. Then

$$X < \frac{1}{\delta}\cdot\log\left(\gamma\cdot C\cdot c/\left(|\tilde{\Lambda}|-\sum_{i=1}^{n}|x_i|\right)\right) . \tag{3.25}$$

Proof. Define the lattice point $\underline{x}$ as in the proof of Lemma 3.7. By (3.23) and (3.24)

$$|\Lambda| \geqslant \left(|\tilde{\Lambda}|-\sum_{i=1}^{n}|x_i|\right)/\gamma\cdot C > 0 .$$

The result follows at once by (3.1). □

We proceed as follows. Choose a constant C_0 such that if $|\tilde{\Lambda}| > C_0$ then the upper bounds for $|x_i|$ imply (3.24). In that case we have a new upper bound for X from (3.25). In case $|\tilde{\Lambda}| \leqslant C_0$ we have an upper bound for the length of the vector $\underline{x}$. We compute all lattice points satisfying this bound by the algorithm of Fincke and Pohst, and check them for (3.1).

Summarizing, the reduction method presented above is based on the fact that a large solution of (3.1) corresponds to an extremely short vector in an appropriate approximation lattice. Since we can actually prove by computations that such short vectors do not exist, it follows that such large solutions do not exist. We will apply these techniques in Chapter 5.

3.8. Inhomogeneous multi-dimensional approximation in the real case: an alternative for the generalized Davenport lemma.

Let Λ be the most general linear form that we will study, viz.

$$\Lambda = \beta + \sum_{i=1}^{n} x_i \cdot \vartheta_i ,$$

where $n \geqslant 2$ (the case $n = 2$ has been dealt with in Section 3.3, but can be incorporated here also). To deal with this inhomogeneous case, two methods are available. The first method is a generalization of the method of Davenport that we discussed in Section 3.3. The second method is closer to the homogeneous case of the previous section.

First we explain briefly the generalized Davenport method. See Ellison [1971[a]] (where only the case $n = 3$ is treated). Put

$$\vartheta'_i = \vartheta_i/\vartheta_n \quad \text{for} \quad i = 1, \ldots, n-1 , \quad \beta' = \beta/\vartheta_n ,$$

$$\Lambda' = \Lambda/\vartheta_n = \beta' + \sum_{i=1}^{n-1} x_i \cdot \vartheta'_i + x_n .$$

Let $(p_1,\ldots,p_{n-1},q)$ be a simultaneous approximation to $\vartheta'_1, \ldots, \vartheta'_{n-1}$ with q of the size of X_0^{n-1} , such that, for $i = 1, \ldots, n-1$,

$$|\vartheta'_i - p_i/q| < c'/q^{1+1/(n-1)}$$

for a small constant c' .

<u>LEMMA 3.9. (Davenport, Ellison).</u> *Suppose that*

$$\|q\cdot\beta'\| > 2\cdot(n-1)\cdot X_0\cdot c'/q^{1/(n-1)} .$$

Then the solutions of (3.1), (3.2) *satisfy*

$$X < \frac{1}{\delta}\cdot\log\left(q^{1+1/(n-1)}\cdot c/|\vartheta_n|\cdot c'\cdot(n-1)\cdot X_0\right) .$$

<u>Proof.</u> The result follows at once from

$$\|q\cdot\beta'\| \leqslant |q\cdot\Lambda' + \sum_{i=1}^{n-1} x_i\cdot(p_i - q\cdot\vartheta'_i)| \leqslant$$

$$q\cdot|\vartheta_n|^{-1}\cdot c\cdot\exp(-\delta\cdot X) + (n-1)\cdot X_0\cdot c'/q^{1/(n-1)} . \qquad \square$$

To apply this generalized Davenport method in practice, it is necessary to compute the simultaneous approximations $(p_1,\dots,p_{n-1},q)$. We indicated in Section 1.4 how this can be done with the L^3-algorithm. As lattice we take the one associated to the following matrix:

$$\begin{pmatrix} 1 & & & \\ [C\cdot\vartheta'_1] & -C & & \emptyset \\ \vdots & & \ddots & \\ [C\cdot\vartheta'_{n-1}] & \emptyset & & -C \end{pmatrix},$$

where C is a constant of size X_0^n . Then $\underline{c}_1$, the first basis vector of a reduced basis, will have length of the size of $C^{(n-1)/n} \cong X_0^{n-1}$. But $\underline{c}_1$ can be written as

$$\underline{c}_1 = \left(q,\ q\cdot[C\cdot\vartheta'_1]-C\cdot p_1,\ \dots,\ q\cdot[C\cdot\vartheta'_{n-1}]-C.p_{n-1} \right)^T$$

for some $p_1, \dots, p_{n-1}, q$. It is expected that q is of size X_0^{n-1} , and

$$q\cdot C\cdot|\vartheta'_i-p_i/q| \cong |q\cdot[C\cdot\vartheta'_i]-C\cdot p_i|$$

are of the size X_0^{n-1} , so that $|\vartheta'_i-p_i/q|$ are of the size

$$X_0^{n-1}/C\cdot X_0^{n-1} = C^{-1} \cong X_0^{-n} \cong q^{-(1+1/(n-1))} ,$$

as desired.

The above method has been applied in practice to solve Thue and Thue-Mahler equations by Agrawal, Coates, Hunt and van der Poorten [1980] (using multi-dimensional continued fractions instead of the L^3-algorithm), Pethö and Schulenberg [1987], and Blass, Glass, Meronk and Steiner [1987^a], [1987^b]. So it has proved to be useful. However, we prefer another method, for several reasons. Firstly, it is close to the homogeneous case as described in the previous section, whereas the generalized Davenport method has no obvious counterpart for the homogeneous case. Secondly, it actually produces solutions for which the linear form Λ is almost as near to zero as possible under the condition $X \leqslant X_0$. Specifically, if a linear relation between the ϑ_i exists, but had not been noticed before (a situation that may occur in practice, cf. Agrawal, Coates, Hunt and van der Poorten [1980]), the method detects these relations, by finding explicitly an extremely short lattice vector (resp. a lattice vector extremely near to a given point) giving the coefficients of the relation. Thirdly, an analogous method for the p-adic case can be given (see Section 3.11). Finally, variations as indicated in Section 1.4 are possible. Concerning computation time we think that the two methods are about equally fast.

The method works as follows. We take the approximation lattice Γ exactly as in the homogeneous case (cf. the previous section), with constants γ, C chosen properly, i.e. C is of the size X_0^n. Compute with the L^3-algorithm a reduced basis $\underline{c}_1, \ldots, \underline{c}_n$ of Γ. Let $\mathcal{C}$ be the matrix associated to this basis, and compute also the transformation matrix $\mathcal{U}$ with $\mathcal{C} = \mathcal{B}\cdot\mathcal{U}$, and its inverse $\mathcal{U}^{-1}$. Note that $\mathcal{B}^{-1}$, and hence also $\mathcal{C}^{-1}$, are easy to compute, namely by

$$\mathcal{B}^{-1} = \begin{pmatrix} 1/\gamma & & & \emptyset \\ & \ddots & & \\ \emptyset & & 1/\gamma & \\ -\dfrac{[\gamma\cdot C\cdot\vartheta_1]}{\gamma\cdot[\gamma\cdot C\cdot\vartheta_n]} & \cdots & -\dfrac{[\gamma\cdot C\cdot\vartheta_{n-1}]}{\gamma\cdot[\gamma\cdot C\cdot\vartheta_n]} & \dfrac{1}{[\gamma\cdot C\cdot\vartheta_n]} \end{pmatrix}$$

and our version of the L^3-algorithm (Fig. 1). Let $\underline{y} \in \mathbb{Z}^n$ be defined by

$$\underline{y} = \big(\,0,\ \ldots,\ 0,\ -[\gamma\cdot C\cdot\beta]\,\big)^T = \sum_{i=1}^{n} s_i\cdot\underline{c}_i\ ,$$

where the coefficients $s_i \in \mathbb{R}$ can be computed by

$$(s_1,\ldots,s_n)^T = \mathcal{C}^{-1}\cdot\underline{y}\ .$$

To be more precise, if $\mathcal{U}^{-1}$ has $\underline{u}$ as n th column, then $\mathcal{C}^{-1}$ has $\underline{u}/[\gamma\cdot C\cdot\vartheta_n]$ as n th column, so

$$(s_1,\ldots,s_n)^T = -\underline{u}\cdot[\gamma\cdot C\cdot\beta]/[\gamma\cdot C\cdot\vartheta_n]\ .$$

Now we apply Lemma 3.5 or 3.6, that provide a lower bound for $\ell(\Gamma,\underline{y})$. Then we can apply the following lemma.

<u>LEMMA 3.10.</u> *Let* X_1 *be a positive constant such that*

$$\ell(\Gamma,\underline{y}) \geq \sqrt{\left((n+2)^2+(n-1)\gamma^2\right)}\cdot X_1\ . \tag{3.26}$$

Then (3.1) *has no solutions with*

$$\frac{1}{\delta}\cdot\log(\gamma\cdot C\cdot c/X_1) \leq X \leq X_1\ . \tag{3.27}$$

<u>Remark.</u> We apply this lemma for $X_1 = X_0$. If condition (3.26) then fails, we must take a larger constant C. If it holds for a constant C of the

size X_0^n , then (3.27) yields a reduced lower bound for X of size $\log X_0$.

Proof. Let $x_1, \ldots, x_n$ be a solution of (3.1) with $0 < X \leqslant X_1$. Consider the lattice point

$$\underline{x} = \sum_{i=1}^{n} x_i \cdot \underline{b}_i = (\gamma \cdot x_1, \ldots, \gamma \cdot x_{n-1}, \tilde{\Lambda}_0)^T ,$$

with

$$\tilde{\Lambda}_0 = \sum_{i=1}^{n} x_i \cdot [\gamma \cdot C \cdot \vartheta_i] .$$

Put $\tilde{\Lambda} = [\gamma \cdot C \cdot \beta] + \tilde{\Lambda}_0$. Then

$$|\underline{x}-\underline{y}|^2 = \gamma^2 \cdot \sum_{i=1}^{n-1} x_i^2 + \tilde{\Lambda}^2 \leqslant (n-1) \cdot \gamma^2 \cdot X_1^2 + \tilde{\Lambda}^2 ,$$

and

$$|\tilde{\Lambda}-\gamma \cdot C \cdot \Lambda| \leqslant |[\gamma \cdot C \cdot \beta]-\gamma \cdot C \cdot \beta| + \sum_{i=1}^{n} |x_i| \cdot |[\gamma \cdot C \cdot \vartheta_i]-\gamma \cdot C \cdot \vartheta_i|$$

$$\leqslant 1 + \sum_{i=1}^{n} |x_i| \leqslant 1 + n \cdot X_1 \leqslant (n+1) \cdot X_1 .$$

By (3.1), (3.26) and the definition of $\ell(\Gamma, \underline{y})$ the result follows, since

$$\gamma \cdot C \cdot c \cdot \exp(-\delta \cdot X) > |\gamma \cdot C \cdot \Lambda| \geqslant |\tilde{\Lambda}| - |\tilde{\Lambda}-\gamma \cdot C \cdot \Lambda|$$

$$\geqslant \sqrt{\left(\ell(\Gamma, \underline{y})^2-(n-1) \cdot \gamma^2 \cdot X_1^2\right)} - (n+1) \cdot X_1 \geqslant X_1 . \qquad \square$$

Again we may prove refinements of the above lemma, similar to Lemma 3.8 in the homogeneous case. We explained in Section 3.5. how to apply the Fincke and Pohst algorithm in the inhomogeneous case. We do not work that out here.

Summarizing, the method described above is based on the fact that a large solution of (3.1) in the inhomogeneous case leads to a lattice point extremely near to a fixed point in $\mathbb{Z}^n$. We can actually prove by some computations that such lattice points do not exist, so that such extreme solutions do not exist. The method outlined in this section is used in Chapter 8. Note that in the case $n = 2$ the method is essentially the same as the Davenport lemma.

3.9. Inhomogeneous zero-dimensional approximation in the p-adic case.

In the p-adic case we start with a very simple linear form Λ , to which also a very simple reduction method applies. Let Λ be

$$\Lambda = \beta + x\cdot\vartheta ,$$

for $\beta, \vartheta \in \Omega_p$ such that $\beta/\vartheta \in \mathbb{Q}_p$, and $x \in \mathbb{Z}$, $x > 0$. It is obvious that in the real case with such a simple linear form Λ inequality (3.1) has only finitely many solutions (we even don't need (3.2)), that are easy to compute. In the p-adic case however, inequality (3.3) may have infinitely many solutions, so we do need a bound like (3.4), and a reduction method.

Put $\vartheta' = -\beta/\vartheta$. Then $\vartheta' \in \mathbb{Q}_p$. Inequality (3.3) now becomes

$$\mathrm{ord}_p(\vartheta'-x) \geqslant c_1' + c_2\cdot x , \tag{3.28}$$

where c_1', c_2 are constants with $c_2 > 0$. We assume that

$$x \geqslant -c_1'/c_2 .$$

Then (3.28) has no solutions if $\mathrm{ord}_p(\vartheta') < 0$. Hence we may assume that ϑ' is a p-adic integer. Let the p-adic expansion of ϑ' be

$$\vartheta' = \sum_{i=0}^{\infty} u_i\cdot p^i ,$$

where $u_i \in \{ 0, 1, ..., p-1 \}$ for all $i \in \mathbb{N}_0$. Compute the p-adic digits u_i far enough to be able to apply the following reduction lemma.

<u>LEMMA 3.11.</u> *Let* X_1 *be a positive constant. Let* r *be the minimal index such that*

$$p^r > X_1 , \quad u_r \neq 0 . \tag{3.29}$$

Then (3.28) *has no solutions with*

$$(r-c_1')/c_2 < x \leqslant X_1 . \tag{3.30}$$

<u>Remark.</u> We apply the lemma with $X_1 = X_0$. The assumption behind the lemma is that in the p-adic expansion of ϑ' no long sequences of zeroes appear. In fact, it seems that in our applications the numbers u_i are distributed randomly over $\{ 0, 1, ..., p-1 \}$. Then the minimal r satisfying (3.29)

will not be much larger than $\log X_0/\log p$, and then (3.30) yields a reduced upper bound of size $\log X_0$, as desired.

Proof. Let $x \leqslant X_1$ satisfy (3.28). Suppose that $\mathrm{ord}_p(\vartheta'-x) \geqslant r+1$. Then

$$x \equiv \sum_{i=0}^{r} u_i \cdot p^i \pmod{p^{r+1}} .$$

By $x \geqslant 0$ it follows from (3.29) that

$$x \geqslant \sum_{i=0}^{r} u_i \cdot p^i \geqslant u_r \cdot p^r \geqslant p^r > X_1 ,$$

which contradicts the assumption $x \leqslant X_1$. Hence $\mathrm{ord}_p(\vartheta'-x) \leqslant r$, and (3.30) follows from (3.28). □

Remark. In the above proof it is essential that $x \geqslant 0$. It is however not difficult to formulate a similar result that holds for all $x \in \mathbb{Z}$, by looking, if $p \neq 2$ for p-adic digits u_i that are not only $\neq 0$ but also $\neq p-1$, and if $p = 2$ for p-adic digits u_i, u_{i+1} with $u_i \neq u_{i+1}$.

A method very similar to the one described above was used by Wagstaff [1979], [1981], a.o. for solving $5^n \equiv 2 \pmod{3^n}$. We apply the method in Chapter 4.

3.10. Homogeneous one-dimensional approximation in the p-adic case: p-adic continued fractions and approximation lattices of p-adic numbers.

Let Λ have the form

$$\Lambda = x_1 \cdot \vartheta_1 + x_2 \cdot \vartheta_2 ,$$

where ϑ_1, $\vartheta_2 \in \Omega_p$ such that $\vartheta = -\vartheta_1/\vartheta_2 \in \mathbb{Q}_p$, and x_1, $x_2 \in \mathbb{Z}$. We may assume that $\mathrm{ord}_p(\vartheta) \geqslant 0$. Now

$$\Lambda' = \Lambda/\vartheta_1 = -x_1 \cdot \vartheta + x_2 .$$

So (3.3) now means that the rational number x_2/x_1 is p-adically close to the p-adic number ϑ .

In analogy of the real case it seems reasonable to study p-adic continued fraction algorithms. However, a p-adic continued fraction algorithm that provides all best approximations to a p-adic number seems not to exist.

Therefore we introduce the concept of p-*adic approximation lattices*, as was done in de Weger [1986[a]]. From this paper we adopt the best approximation algorithm, which is a generalization of the algorithm of Mahler [1961], Chapter IV. This algorithm goes back also on the euclidean algorithm, and thus is close to a continued fraction algorithm. But it is not a p-adic continued fraction algorithm in the sense that a p-adic number is expanded into a continued fraction, and that the approximations are then found by truncating the continued fraction.

Recall that for $\mu \in \mathbb{N}_0$ the rational integer $\vartheta^{(\mu)}$ is defined by $\mathrm{ord}_p(\vartheta-\vartheta^{(\mu)}) \geq \mu$ and $0 \leq \vartheta^{(\mu)} < p^{\mu}$. We define for any $\mu \in \mathbb{N}_0$ the p-adic approximation lattice Γ_μ by a matrix to which a basis of Γ_μ is associated, namely the matrix

$$\begin{pmatrix} 1 & 0 \\ \vartheta^{(\mu)} & p^{\mu} \end{pmatrix} .$$

Then it is easy to see that

$$\Gamma_\mu = \langle\ (x_1,x_2)^T \in \mathbb{Z}^2 \mid \mathrm{ord}_p(x_2-x_1\cdot\vartheta) \geq \mu\ \rangle$$

(cf. Lemma 3.13 in the next section, where we prove a more general result). The following algorithm computes a point of minimal length in Γ_μ.

```
     x := (1,ϑ^(μ))^T ; y := (0,p^μ)^T ;
     if |x| > |y| , interchange x and y ;
(1)  compute K ∈ ℤ such that |y-K·x| is minimal ;
     y := y - K·x ;
     if |x| > |y| , interchange x and y , and go to (1) ;
     print x .
```

Figure 3. p-adic approximation algorithm.

With this algorithm it is possible to compute $\ell(\Gamma_\mu)$ explicitly. Then we can apply the following lemma.

LEMMA 3.12. *Let* X_1 *be a constant such that*

$$\ell(\Gamma_\mu) > \sqrt{2}\cdot X_1 . \tag{3.31}$$

Then (3.3) *has no solutions with*

$$\left(\mu-1-c_1+\mathrm{ord}_p(\vartheta_2)\right)/c_2 < x_j \leqslant X \leqslant X_1 . \tag{3.32}$$

<u>Remark.</u> We take μ such that p^{μ} is of the size of X_0^2, and apply the lemma for $X_1 = X_0$. Then we expect that $\ell(\Gamma_{\mu})$ is of the size of X_0, so that (3.31) is a reasonable condition.

<u>Proof.</u> Apply the proof of Lemma 3.14 (in the next section) for $n = 2$. □

A method like the one described above has been applied by Agrawal, Coates, Hunt and van der Poorten [1980]. We use it in Chapters 6 and 7.

3.11. Homogeneous multi-dimensional approximation in the p-adic case: p-adic approximation lattices.

We now study the case

$$\Lambda = \sum_{i=1}^{n} x_i \cdot \vartheta_i ,$$

where $\vartheta_i \in \Omega_p$ such that $\vartheta_i/\vartheta_j \in \mathbb{Q}_p$, $x_i \in \mathbb{Z}$ for all i, j, and with $n \geqslant 2$. We may assume that $\mathrm{ord}_p(\vartheta_i)$ is minimal for $i = n$. Put

$$\vartheta'_i = -\vartheta_i/\vartheta_n \quad \text{for} \quad i = 1, \ldots, n-1 .$$

Then $\vartheta'_i \in \mathbb{Z}_p$ for all i. Put

$$\Lambda' = \Lambda/\vartheta_n = -\sum_{i=1}^{n-1} x_i \cdot \vartheta'_i + x_n .$$

The definition of the p-adic approximation lattices can be generalized directly from the one-dimensional case. Namely, for any $\mu \in \mathbb{N}_0$ we define Γ_{μ} as the lattice associated to the matrix

$$\mathcal{B}_{\mu} = \begin{pmatrix} 1 & & & \emptyset \\ & \ddots & & \\ \emptyset & & 1 & \\ \vartheta_1'^{(\mu)} & \ldots & \vartheta_{n-1}'^{(\mu)} & p^{\mu} \end{pmatrix} .$$

Then we have the following result.

LEMMA 3.13. *The lattice* Γ_μ *, associated to the above defined matrix* $\mathcal{B}_\mu$ *, is equal to the set*

$$\Gamma_\mu = \langle \, (x_1,\ldots,x_n)^T \in \mathbb{Z}^n \mid \mathrm{ord}_p(\Lambda') \geq \mu \, \rangle \, .$$

Proof. For any $\underline{x} = (x_1,\ldots,x_n)^T \in \Gamma_\mu$ there exists a $\underline{z} = (z_1,\ldots,z_n)^T \in \mathbb{Z}^n$ such that $\underline{x} = \mathcal{B}_\mu \cdot \underline{z}$. Then $x_i = z_i$ for $i = 1, \ldots, n-1$, and

$$x_n = \sum_{i=1}^{n-1} z_i \cdot \vartheta_i'^{(\mu)} + z_n \cdot p^\mu \equiv \sum_{i=1}^{n-1} x_i \cdot \vartheta'_i \pmod{p^\mu} \, .$$

Hence $\mathrm{ord}_p(\Lambda') \geq \mu$. Conversely, for any $\underline{x} = (x_1,\ldots,x_n)^T$ such that $\mathrm{ord}_p(\Lambda') \geq \mu$ there obviously exists a $\underline{z} \in \mathbb{Z}^n$ such that $\underline{x} = \mathcal{B}_\mu \cdot \underline{z}$. □

Using the L^3-algorithm we can compute a lower bound for $\ell(\Gamma_\mu)$. Then we can apply the following lemma, which is a direct generalization of Lemma 3.12.

LEMMA 3.14. *Let* X_1 *be a constant such that*

$$\ell(\Gamma_\mu) > \sqrt{n} \cdot X_1 \, . \tag{3.33}$$

Then (3.3) *has no solutions with*

$$\big(\mu - 1 - c_1 + \mathrm{ord}_p(\vartheta_n)\big)/c_2 < x_j \leq X \leq X_1 \, . \tag{3.34}$$

Remark. We take μ such that p^μ is of the size of X_0^n , and apply the lemma for $X_1 = X_0$. Then we expect that $\ell(\Gamma_\mu)$ is of the size of X_0 , so that (3.33) is a reasonable condition.

Proof. Let $x_1, \ldots, x_n$ be a solution of (3.3) with $X \leq X_1$. Then (3.33) prohibits the point $(x_1,\ldots,x_n)^T$ from being a lattice point in Γ_μ . Hence, by Lemma 3.13, $\mathrm{ord}_p(\Lambda') \leq \mu-1$, and (3.34) follows from (3.3). □

We will apply the results of this section in Chapters 6 and 7.

3.12. Inhomogeneous one- and multi-dimensional approximation in the p-adic case.

Finally we study an inhomogeneous p-adic form

$$\Lambda = \beta + \sum_{i=1}^{n} x_i \cdot \vartheta_i ,$$

where $\beta, \vartheta_i \in \Omega_p$ such that $\beta/\vartheta_j, \vartheta_i/\vartheta_j \in \mathbb{Q}_p$ and $x_i \in \mathbb{Z}$ for all i, j, and $n \geqslant 2$. We assume that $\mathrm{ord}_p(\vartheta_i)$ is minimal for $i = n$, and that $\mathrm{ord}_p(\beta) \geqslant \mathrm{ord}_p(\vartheta_n)$. Put

$$\vartheta'_i = -\vartheta_i/\vartheta_n \quad \text{for} \quad i = 1, \ldots, n-1 , \quad \beta' = \beta/\vartheta_n ,$$

$$\Lambda' = \Lambda/\vartheta_n = \beta' - \sum_{i=1}^{n-1} x_i \cdot \vartheta'_i + x_n .$$

Then $\beta', \vartheta'_i \in \mathbb{Z}_p$ for all i. As p-adic approximation lattices we take the lattices Γ_μ that were defined for the homogeneous case, i.e. for any $\mu \in \mathbb{N}_0$ the lattice Γ_μ that is associated to the matrix $\mathcal{B}_\mu$ (see Section 3.11). Further put

$$\underline{y} = \left(0, \ldots, 0, \beta'^{(\mu)} \right)^T = \sum_{i=1}^{n} s_i \cdot \underline{c}_i \in \mathbb{Z}^n ,$$

where $\underline{c}_1, \ldots, \underline{c}_n$ is a reduced basis of Γ_μ, and $s_i \in \mathbb{R}$. By Lemma 3.5 or 3.6 we can compute a lower bound for $\ell(\Gamma,\underline{y})$. This is useful in view of the following lemma.

<u>LEMMA 3.15.</u> *The set* $\Gamma_\mu(\underline{y}) = \Gamma_\mu + \underline{y}$ *is equal to the set*

$$\Gamma_\mu(\underline{y}) = \langle \left(x_1,\ldots,x_n\right)^T \in \mathbb{Z}^n \mid \mathrm{ord}_p(\Lambda') \geqslant \mu \rangle .$$

<u>Proof.</u> Let $\underline{x} = \left(x_1,\ldots,x_n\right)^T$ satisfy $\underline{x} - \underline{y} \in \Gamma_\mu$. Note that

$$\underline{x} - \underline{y} = \left(x_1, \ldots, x_{n-1}, x_n - \beta'^{(\mu)} \right)^T .$$

By Lemma 3.13 we have

$$\mathrm{ord}_p\Big(\sum_{i=1}^{n-1} x_i \cdot \vartheta_i - (x_n - \beta'^{(\mu)})\Big) \geqslant p^\mu .$$

The left hand side is just $\mathrm{ord}_p(\Lambda')$, which proves the lemma. □

Obviously, the length of the shortest vector in $\Gamma_\mu(\underline{y})$ (a translated lattice) is equal to $\ell(\Gamma_\mu,\underline{y})$ (unless in the case $\underline{y} \in \Gamma_\mu$, i.e. $s_i \in \mathbb{Z}$ for all i). We have the following useful lemma.

LEMMA 3.16. *Let* X_1 *be a constant such that*

$$\ell(\Gamma_\mu, \underline{y}) > \sqrt{n}\cdot X_1 . \tag{3.35}$$

Then (3.3) *has no solutions with*

$$(\mu-1-c_1+\mathrm{ord}_p(\vartheta_n))/c_2 < x_j \leq X \leq X_1 . \tag{3.36}$$

Remark. We take μ such that p^μ is of the size of X_0^n , and apply the lemma for $X_0 = X_1$. Then we expect that $\ell(\Gamma_\mu, \underline{y})$ is of the size of X_0 , so that (3.35) is a reasonable condition.

Proof. Let $x_1, \ldots, x_n$ be a solution of (3.3) with $X \leq X_1$. Then (3.35) prohibits the point $(x_1,\ldots,x_n)^T$ from being in $\Gamma_\mu(\underline{y})$. Hence, by Lemma 3.15, $\mathrm{ord}_p(\Lambda') \leq \mu-1$, and (3.36) follows from (3.3). □

We will not apply the above lemma in this book. It is included here only for the sake of completeness. However, when solving Thue-Mahler equations (see Section 8.6), it will be of use.

3.13. Useful sublattices of p-adic approximation lattices.

In our p-adic applications of solving diophantine equations via linear forms, we always have linear forms in logarithms of algebraic numbers, i.e. in

$$\Lambda = \beta + \sum_{i=1}^{n} x_i\cdot\vartheta_i$$

the β and ϑ_i's are p-adic logarithms of algebraic numbers, say

$$\beta = \log_p(\alpha_0) , \quad \vartheta_i = \log_p(\alpha_i) \quad \text{for } i = 1, \ldots, n .$$

In Section 2.3 we have seen that for a $\xi \in \mathbb{Q}_p$ if $\mathrm{ord}_p(1\pm\xi) > 1/(p-1)$ then $\mathrm{ord}_p(\log_p(\xi)) = \mathrm{ord}_p(1\pm\xi)$. In our applications we apply this to

$$\xi = \alpha_0\cdot\prod_{i=1}^{n}\alpha_i^{x_i} ,$$

for which $\mathrm{ord}_p(\xi-1)$ is large. This implies that $\mathrm{ord}_p(\log_p(\xi))$ is large too, on which we based the definition of our approximation lattices. However, the converse is not necessarily true: $\mathrm{ord}_p(\log_p(\xi))$ being large does not imply that $\mathrm{ord}_p(\xi-1)$ is large. This is due to the fact that the p-adic logarithm is a multi-branched function. To be more precise, for any root of unity $\zeta \in \mathbb{Q}_p$ we have $\log_p(\zeta) = 0$ (cf. Section 2.3). In $\mathbb{Q}_p$ there exist only the $(p-1)$ th roots of unity if p is odd, and only ± 1 as roots of unity if $p = 2$. Let ζ be a primitive $(p-1)$ th root of unity if p is odd, and $\zeta = -1$ if $p = 2$. It follows that $\mathrm{ord}_p(\log_p(\xi))$ being large implies that for some $k \in \{ 0, 1, \ldots, p-2 \}$ (or $k \in \{ 0, 1 \}$ if $p = 2$)

$$\mathrm{ord}_p(\log_p(\xi)) = \mathrm{ord}_p(\xi - \zeta^k) .$$

The set of $x_1, \ldots, x_n$ such that $\mathrm{ord}_p(\xi-1)$ (or $\mathrm{ord}_p(\xi\pm1)$ if one wishes) is large, turns out to be a sublattice Γ^*_μ (or $\Gamma^\#_\mu$ respectively) of Γ_μ . In the following lemma we shall prove this fact, and indicate how a basis of such a sublattice can be found. Then we can work with this sublattice instead of Γ_μ itself. Of course, in Lemmas 3.12, 3.14 and 3.16 we can replace Γ_μ by these sublattices Γ^*_μ, $\Gamma^\#_\mu$. For simplicity we assume that $\alpha_i \in \mathbb{Q}_p$ for all i . We take $\alpha_0 = 1$ (corresponding to $\beta = 0$, thus to the homogeneous case), and leave it to the reader to define appropriate translated lattices $\Gamma^*_\mu(\underline{y})$, $\Gamma^\#_\mu(\underline{y})$ for the case $\alpha_0 \neq 1$ (the inhomogeneous case).

<u>LEMMA 3.17.</u> (i). *Let* $\alpha_1, \ldots, \alpha_n \in \mathbb{Q}_p$ *be given numbers with* $\mathrm{ord}_p(\alpha_i) = 0$ *for all* i , *and* $\mathrm{ord}_p(\log_p(\alpha_i))$ *minimal for* $i = n$. *Let* $x_1, \ldots, x_n \in \mathbb{Z}$. *Put*

$$\xi = \prod_{i=1}^{n} \alpha_i^{x_i} , \quad \mu_0 = \mathrm{ord}_p(\log_p(\alpha_n)) .$$

For any $\mu \in \mathbb{N}_0$ *put*

$$\Gamma_\mu = \langle\, (x_1,\ldots,x_n) \in \mathbb{Z}^n \mid \mathrm{ord}_p(\log_p(\xi)) \geqslant \mu + \mu_0 \,\rangle ,$$

$$\Gamma^*_\mu = \langle\, (x_1,\ldots,x_n) \in \mathbb{Z}^n \mid \mathrm{ord}_p(\xi\pm1) \geqslant \mu + \mu_0 \,\rangle ,$$

$$\Gamma^\#_\mu = \langle\, (x_1,\ldots,x_n) \in \mathbb{Z}^n \mid \mathrm{ord}_p(\xi-1) \geqslant \mu + \mu_0 \,\rangle .$$

Then $\Gamma^\#_\mu \subseteq \Gamma^*_\mu \subseteq \Gamma_\mu$ *are lattices. If* $p = 2$ *they are all equal. If* $p = 3$ *then* $\Gamma^*_\mu = \Gamma_\mu$. *If* $p \geqslant 3$ *then* $\#(\Gamma_\mu/\Gamma^*_\mu) = (p-1)/2$, $\#(\Gamma_\mu/\Gamma^\#_\mu) = p-1$, $\#(\Gamma^*_\mu/\Gamma^\#_\mu) = 2$.
(ii). *Let* $\underline{b}_1, \ldots, \underline{b}_n$ *be a basis of* Γ_μ . *Define* $k(\underline{x})$ *for any* $\underline{x} = (x_1,\ldots,x_n)^T \in \Gamma_\mu$ *by*

$$\xi \equiv \zeta^{k(\underline{x})} \pmod{p^{\mu+\mu_0}} , \quad k(\underline{x}) \in \{ 0, 1, \ldots, p-2 \} .$$

Let $\underline{b}'_1, \ldots, \underline{b}'_n$ *be a basis of* Γ_μ *such that*

$$k(\underline{b}'_n) = \gcd\big(k(\underline{b}_1),\ldots,k(\underline{b}_n)\big) .$$

Put for $i = 1, \ldots, n-1$ *and* $p \geqslant 5$

$$\gamma^*_i \equiv k(\underline{b}'_i)/k(\underline{b}'_n) \pmod{(p-1)/2} , \quad |\gamma^*_i| \leqslant (p-1)/4 ,$$

$$\underline{b}^*_i = \underline{b}'_i - \gamma^*_i\cdot\underline{b}'_n ,$$

and for $p \geqslant 3$ *also*

$$\gamma^\#_i \equiv k(\underline{b}'_i)/k(\underline{b}'_n) \pmod{(p-1)} , \quad |\gamma^\#_i| \leqslant (p-1)/2 ,$$

$$\underline{b}^\#_i = \underline{b}'_i - \gamma^\#_i\cdot\underline{b}'_n .$$

Further put for $p \geqslant 5$

$$\gamma^*_n = \mathrm{lcm}\big(k(\underline{b}'_n),(p-1)/2\big)/k(\underline{b}'_n) , \quad \underline{b}^*_n = \gamma^*_n\cdot\underline{b}'_n ,$$

and for $p \geqslant 3$ *also*

$$\gamma^\#_n = \mathrm{lcm}\big(k(\underline{b}'_n),p-1\big)/k(\underline{b}'_n) , \quad \underline{b}^\#_n = \gamma^\#_n\cdot\underline{b}'_n .$$

Then $\underline{b}^*_1, \ldots, \underline{b}^*_n$ *is a basis of* Γ^*_μ *, and* $\underline{b}^\#_1, \ldots, \underline{b}^\#_n$ *is a basis of* $\Gamma^\#_\mu$ *.*

<u>Proof.</u> (i). It is trivial that $\Gamma^\#_\mu \subseteq \Gamma^*_\mu \subseteq \Gamma_\mu$, and that they are lattices. The equalities of the lattices for $p = 2, 3$ follow from the fact that ± 1 are the only roots of unity in $\mathbb{Q}_p$ for $p = 2, 3$. The values of $\#(\Gamma_\mu/\Gamma^*_\mu)$, etc., follow from (ii).

(ii). Note that $k(\underline{x})$ is (mod (p-1)) a linear function on Γ_μ . The points $\underline{x}$ of Γ^*_μ are characterized by $(p-1)/2 \mid k(\underline{x})$, and the points $\underline{x}$ of $\Gamma^\#_m$ are characterized by $(p-1) \mid k(\underline{x})$. It follows from the definitions in the lemma that for $i = 1, \ldots, n-1$

$$k(\underline{b}^*_i) \equiv k(\underline{b}'_i) - \gamma^*_i\cdot k(\underline{b}'_n) \equiv 0 \pmod{(p-1)/2} ,$$

$$k(\underline{b}^\#_i) \equiv k(\underline{b}'_i) - \gamma^\#_i\cdot k(\underline{b}'_n) \equiv 0 \pmod{(p-1)} .$$

Note that $\underline{b}^*_1, \ldots, \underline{b}^*_{n-1}, \underline{b}'_n$ and $\underline{b}^\#_1, \ldots, \underline{b}^\#_{n-1}, \underline{b}'_n$ are both bases of Γ_μ . Write $\underline{x} \in \Gamma_\mu$ as

$$\underline{x} = \sum_{i=1}^{n-1} y^*_i\cdot\underline{b}^*_i + y^*_n\cdot\underline{b}'_n = \sum_{i=1}^{n-1} y^\#_i\cdot\underline{b}^\#_i + y^\#_n\cdot\underline{b}'_n$$

for integers y^*_i, $y^\#_i$. Then it follows that

$$k(\underline{x}) \equiv y_n^{*} \cdot k(\underline{b}'_n) \pmod{(p-1)/2} ,$$

$$k(\underline{x}) \equiv y_n^{\#} \cdot k(\underline{b}'_n) \pmod{(p-1)} .$$

So $\underline{x} \in \Gamma_\mu^{*}$ if and only if $\gamma_n^{*} \mid y_n^{*}$, and $\underline{x} \in \Gamma_\mu^{\#}$ if and only if $\gamma_n^{\#} \mid y_n^{\#}$. This proves the result. $\square$

Chapter 4. S-integral elements of binary recurrence sequences.

Acknowledgements. The research for this chapter has been done partly in cooperation with A. Pethö from Debrecen. The results have been published in Pethö and de Weger [1986] and de Weger [1986[b]].

4.1. Introduction.

In this chapter we present a reduction algorithm for the following problem. Let A, B, G_0, G_1 be integers, and let the recurrence sequence $\{G_n\}_{n=0}^{\infty}$ be defined by

$$G_{n+1} = A \cdot G_n - B \cdot G_{n-1} \quad \text{for} \quad n = 1, 2, \ldots .$$

Assume that $\Delta = A^2 - 4 \cdot B$ is not a square, and that the sequence is not degenerate (this will be explained below). Let w be a nonzero integer, and let $p_1, \ldots, p_s$ be distinct primes. We study the diophantine equation

$$G_n = w \cdot \prod_{i=1}^{s} p_i^{m_i} \tag{4.1}$$

in nonnegative integers $n, m_1, \ldots, m_s$. We will study both the cases of positive and negative discriminant Δ (the 'hyperbolic' and 'elliptic' cases). It was shown by Mahler [1934] that (4.1) has only finitely many solutions. For the case $\Delta > 0$ Schinzel [1967] has given an effectively computable upper bound for the solutions.

Mignotte [1984[a]], [1984[b]] indicated how in some instances (4.1) with $s = 1$ can be solved by congruence techniques. It is however not clear that his method will work for any equation (4.1) with $s = 1$. Moreover, his method seems not to be generalizable for $s > 1$. Pethö [1985] has given a reduction algorithm, based on the Gelfond-Baker method, to treat (4.1) in the case $\Delta > 0$, $w = s = 1$.

Our reduction algorithms are based on a simple case of p-adic diophantine approximation, namely the zero-dimensional case, cf. Section 3.9. In the hyperbolic case this suffices to be able to find all solutions of (4.1). This is based on a trivial observation on the exponential growth of $|G_n|$ in this case. In the elliptic case the situation is essentially more complicated. Then information on the growth of $|G_n|$ can be obtained from the complex Gelfond-Baker theory. Therefore in this case we have to combine the p-adic arguments with the one-dimensional homogeneous or inhomogeneous real diophantine approximation method, cf. Sections 3.2 and 3.3.

We shall give explicit upper bounds for the solutions of (4.1) which are small enough to admit the practical application of the reduction algorithms, if the parameters of the equation are not too large. Pethö [1985] pointed out that essentially better upper bounds hold for all but possibly one solutions. His reasoning is essentially the same as our reduction technique.

The generalized Ramanujan-Nagell equation

$$x^2 + k = \prod_{i=1}^{s} p_i^{z_i} , \tag{4.2}$$

where $k \in \mathbb{Z}$ is fixed, and $x, z_1, \ldots, z_s \in \mathbb{N}_0$ are the unknowns, can be reduced to a finite number of equations of type (4.1) with $\Delta > 0$. Equation (4.2) with $s = 1$ has a long history (cf. Hasse [1966], Beukers [1981] for a survey), and interesting applications in coding theory (cf. Bremner, Calderbank, Hanlon, Morton and Wolfskill [1983], MacWilliams and Sloane [1977], and Tzanakis and Wolfskill [1986], [1987]). Examples of (4.2) have been solved using the Gelfond-Baker theory by Hunt and van der Poorten (unpublished). They used real or complex, not p-adic linear forms in logarithms. As far as we know, none of the proposed methods to treat (4.2) gives rise to an algorithm which works for arbitrary values of k and the p_i's , whereas Tzanakis' elementary method (cf. Tzanakis [1983]) seems to be the only one that can be generalized to $s > 1$. Our method has both properties.

This chapter is organized as follows. In Section 4.2 we give some preliminaries on binary recurrence sequences. In Section 4.3 we study the growth of $|G_n|$, both in the hyperbolic and the elliptic case. The hyperbolic case is trivial, and in the elliptic case we give a method for solving $|G_n| < v$ for a fixed $v \in \mathbb{R}$, by proving an upper bound for n that has particularly good dependence on v , and by showing how to reduce such a bound. Section 4.4 gives upper bounds for the solutions of (4.1).

Section 4.5 gives a lemma on which the p-adic part of the reduction procedure is based. Then Section 4.6 treats some special cases, a.o. the 'symmetric' recurrences. For this special type of recurrence sequences our reduction algorithms fail, but elementary arguments will always work for solving (4.1) in these cases. In Section 4.7 we give the algorithm for reducing upper bounds for the solutions of (4.1) in the case $\Delta > 0$, with some elaborated examples. The same is done for the case $\Delta < 0$ in Section 4.8.

Section 4.9 shows how to treat the generalized Ramanujan-Nagell equation (4.2), as an application of the hyperbolic case of (4.1). As an example we determine all integers x such that $x^2 + 7$ has no prime factors larger than 20, thus extending the result of Nagell [1948] on the equation $x^2 + 7 = 2^n$ (the original Ramanujan-Nagell equation). Finally in Section 4.10 we give an application of the elliptic case of (4.1) to a certain type of mixed quadratic-exponential diophantine equation, analogous to the application of the hyperbolic case to solving (4.2). As an example, we determine the solutions X, m_1, m_2, n of

$$X^2 - 3^{m_1}\cdot 7^{m_2}\cdot X + 2\cdot\left(3^{m_1}\cdot 7^{m_2}\right)^2 = 11\cdot 2^n .$$

4.2. Binary recurrence sequences.

Let $A, B, G_0, G_1 \in \mathbb{Z}$ be given. Let the sequence $\{G_n\}_{n=0}^{\infty}$ be defined by

$$G_{n+1} = A\cdot G_n - B\cdot G_{n-1} \quad \text{for } n = 1, 2, \ldots . \tag{4.3}$$

Let α, β be the roots of $x^2 - A\cdot x + B = 0$. We assume that $\Delta = A^2 - 4\cdot B$ is not a square, and that α/β is not a root of unity (i.e. the sequence is not degenerate). Put

$$\lambda = \frac{G_1 - G_0\cdot\beta}{\alpha - \beta} , \quad \mu = \frac{G_0\cdot\alpha - G_1}{\alpha - \beta} . \tag{4.4}$$

Then λ and μ are conjugates in $K = \mathbb{Q}(\sqrt{\Delta})$. We now have for all $n \geqslant 0$

$$G_n = \lambda\cdot\alpha^n + \mu\cdot\beta^n , \tag{4.5}$$

(cf. Shorey and Tijdeman [1986], Theorem C.1). We will show that when we are solving (4.1), we may assume without loss of generality that

$$(G_0,G_1) = (G_1,B) = (A,B) = 1 .$$

Namely, if $d = (G_0,G_1)$ then $d \mid G_n$ for all $n \geqslant 0$, and thus we may study (4.1) with $G'_n = G_n / d$ instead of with G_n . Next suppose that $d = (A,B)$. If also $d^2 \mid B$ then it is easy to show that $d^{n-1} \mid G_n$ for all $n \geqslant 2$. Then we study (4.1) with $G'_n = G_{n+1} / d^n$ instead of with G_n . The A', B' such that $G'_{n+1} = A' \cdot G'_n - B' \cdot G'_{n-1}$ are $A' = A / d$, $B' = B / d^2$, and thus $(A',B') = 1$. If however $d^2 \nmid B$, then we split the sequence into two parts. We study (4.1) first with $G'_n = G_{2 \cdot n}$ and then with $G'_n = G_{2 \cdot n+1}$, instead of with G_n . For both sequences $\{G'_n\}$ the A', B' such that $G'_{n+1} = A' \cdot G'_n - B' \cdot G'_{n-1}$ are given by $A' = A^2 - 2 \cdot B$, $B' = B^2$. Then $(A',B') = d$, and $d^2 \mid B'$, so we are in the previous case. Finally, let p be a prime such that $p \mid (G_1,B)$, and let $\mathfrak{p}$ be a prime ideal of $\mathbb{Q}(\sqrt{\Delta})$ lying above p . By $p \mid B = \alpha \cdot \beta$ we have $\mathfrak{p} \mid (\alpha)$ or $\mathfrak{p} \mid (\beta)$. Suppose $\mathfrak{p} \mid (\alpha)$. Then $\mathfrak{p} \nmid (\beta)$ by $(A,B) = 1$ (note that $A = \alpha + \beta$). Hence

$$\mathrm{ord}_{\mathfrak{p}}(\lambda \cdot \alpha^n + \mu \cdot \beta^n) = \min [\![\mathrm{ord}_{\mathfrak{p}}(\lambda \cdot \alpha^n),\ \mathrm{ord}_{\mathfrak{p}}(\mu \cdot \beta^n)]\!] = \mathrm{ord}_{\mathfrak{p}}(\mu)$$

if $n \geqslant n_0$ for some n_0 . Thus $\mathrm{ord}_p(G_n)$ is constant for $n \geqslant n_0$, and the same is true if $\mathfrak{p} \mid (\beta)$. Thus we may assume that $(G_1,B) = 1$.

<u>LEMMA 4.1.</u> *Let* $n, m_1, \ldots, m_s$ *be a solution of* (4.1). *Then, with the above assumptions, we have for* $i = 1, \ldots, s$ *either* $m_i = 0$ *or* $n = 0$ *or*

$$\begin{aligned} &\mathrm{ord}_{p_i}(\alpha) = \mathrm{ord}_{p_i}(\beta) = 0 , \\ &\mathrm{ord}_{p_i}(\lambda) = \mathrm{ord}_{p_i}(\mu) = -\tfrac{1}{2} \cdot \mathrm{ord}_{p_i}(\Delta) \leqslant 0 . \end{aligned} \tag{4.6}$$

<u>Proof.</u> Suppose $p_i \mid B$. Then $p_i \nmid A$, hence, from (4.3) and $(B,G_1) = 1$, $p_i \nmid G_n$ for all $n \geqslant 1$. Thus, $m_i = 0$ or $n = 0$. Next suppose $p_i \nmid B$. Then, by $\alpha \cdot \beta = B$,

$$\mathrm{ord}_{p_i}(\alpha) + \mathrm{ord}_{p_i}(\beta) = \mathrm{ord}_{p_i}(B) = 0 .$$

Now, α and β are algebraic integers, so their p_i-adic orders are nonnegative. It follows that they are zero. Put $E = -\lambda \cdot \mu \cdot \Delta$. Note that $E \in \mathbb{Z}$, and for all $n \geqslant 0$

$$G_{n+1}^2 - A \cdot G_n \cdot G_{n+1} + B \cdot G_n^2 = E \cdot B^n .$$

Suppose that $p_i \mid E$, then we infer that $p_i \nmid G_n$ for all n, since $(G_0, G_1) = 1$. Hence $m_i = 0$. Next suppose $p_i \nmid E$, then

$$\mathrm{ord}_{p_i}(\lambda\cdot\sqrt{\Delta}) + \mathrm{ord}_{p_i}(\mu\cdot\sqrt{\Delta}) = \mathrm{ord}_{p_i}(E) = 0 .$$

Since $\lambda\cdot\sqrt{\Delta}$ and $\mu\cdot\sqrt{\Delta}$ are algebraic integers (note that $\sqrt{\Delta} = \alpha - \beta$), the result follows. □

From Lemma 2.1 it follows that we may assume without loss of generality that (4.6) holds for $i = 1, \ldots, s$. We may also assume that $\mathrm{ord}_{p_i}(w) = 0$ for $i = 1, \ldots, s$. The special case $s = 0$ in equation (4.1) is trivial if $\Delta > 0$, and will be treated implicitly in the next section for all Δ.

4.3. The growth of the recurrence sequence.

First we treat the hyperbolic case $\Delta > 0$. Note that $|\alpha| \neq |\beta|$, since the sequence is not degenerate. So we may assume $|\alpha| > |\beta|$. We have the following, almost trivial, result on the exponentiality of the growth of the sequence $\{G_n\}_{n=0}^{\infty}$. Let

$$n_0 > \max\left(2, \log\left|\frac{\mu}{\lambda}\right| / \log\left|\frac{\alpha}{\beta}\right|\right),$$

$$\gamma = |\lambda| - |\mu|\cdot\left|\frac{\alpha}{\beta}\right|^{-n_0} .$$

Note that $\gamma > 0$.

<u>LEMMA 4.2.</u> *Let* $\Delta > 0$. *If* $n \geqslant n_0$ *then* $|G_n| \geqslant \gamma\cdot|\alpha|^n$.

<u>Proof.</u> By (4.5), $|\alpha| > |\beta|$ and $n_0 > 0$ it follows for $n \geqslant n_0$ that

$$|G_n|\cdot|\alpha|^{-n} = \left|\lambda+\mu\cdot\left(\frac{\alpha}{\beta}\right)^{-n}\right| \geqslant |\lambda| - |\mu|\cdot\left|\frac{\alpha}{\beta}\right|^{-n} \geqslant \gamma . \qquad \square$$

We apply this to (4.1) as follows.

<u>COROLLARY 4.3.</u> *Let* $\Delta > 0$. *Any solution* $n, m_1, \ldots, m_s$ *of* (4.1) *with* $n \geqslant n_0$ *satisfies*

$$n < \sum_{i=1}^{s} m_i\cdot\frac{\log p_i}{\log|\alpha|} - \frac{\log(\gamma/|w|)}{\log|\alpha|} .$$

Proof. Clear, from Lemma 4.2 and (4.1). □

Next we study the elliptic case $\Delta < 0$. Since α/β is not a root of unity, $B \geqslant 2$. Since (α,β) and (λ,μ) are pairs of complex conjugates, $|\alpha| = |\beta|$ and $|\lambda| = |\mu|$. Let $v \in \mathbb{R}$, $v \geqslant 1$ be given. We study the inequality

$$|G_n| \leqslant v \tag{4.7}$$

in the variable $n \in \mathbb{N}_0$. We apply a result of Waldschmidt (see Section 2.3) from the complex theory of linear forms in logarithms, which gives an upper bound for n that is particularly good in v . See also Kiss [1979]. Let

$$E = -\lambda\cdot\mu\cdot\Delta ,$$

$$U_2 = \tfrac{1}{2}\cdot\max\,(\,\pi,\ \log B\,)\ ,\quad U_3 = \tfrac{1}{2}\cdot\max\,(\,\pi,\ \log E\,)\ ,$$

$$U_2^+ = \min\,(\,U_2,\ U_3\,)\ ,\quad U_3^+ = \max\,(\,U_2,\ U_3\,)\ ,$$

$$C_1 = 3.362\times10^{21}\cdot U_2\cdot U_3\cdot\log(2\cdot e\cdot U_2^+)\ ,\quad C_2 = \log(4\cdot e\cdot U_3^+)\ ,$$

$$C_3 = \max\Big(\log(\pi/2\cdot|\mu|) + C_1\cdot C_2 + C_1\cdot\log(4\cdot C_1/\log B),\ \tfrac{1}{2}\cdot\log|\lambda\cdot\sqrt{\Delta}|\Big)\cdot 4/\log B\ .$$

THEOREM 4.4. *Let* $\Delta < 0,\ v \in \mathbb{R},\ v \geqslant 1$. *If* $n \geqslant 0$ *satisfies* (4.7) *then*

$$n < C_3 + \frac{4}{\log B}\cdot\log v\ .$$

Remark. Note that C_3 does not depend on v .

The following corollary of Theorem 4.4 is immediate.

COROLLARY 4.5. *Let* $\Delta < 0$. *Any solution* $n, m_1, \ldots, m_s$ *of* (4.1) *satisfies*

$$n < C_3 + \frac{4}{\log B}\cdot[\![\ \log|w| + \sum_{i=1}^{s} m_i\cdot\log p_i\ ]\!]\ .$$

Proof (of theorem 4.4). Note that $|\alpha| = |\beta| = \sqrt{B} \geqslant \sqrt{2}$. First we treat the case $G_n = 0$. Kiss [1979] gives an upper bound for such n , but since in our situation $(G_0,G_1) = (G_1,B) = (A,B) = 1$, we can do much better. Namely, put $R_n = (\alpha^n-\beta^n)/(\alpha-\beta)$ for all $n \in \mathbb{Z}$. It is easy to show that $R_n \in \mathbb{Z}$

and $R_{-n} = -B^{-n}\cdot R_n$ for all $n \in \mathbb{Z}$. Now $G_{n_0} = \lambda\cdot\alpha^{n_0} + \mu\cdot\beta^{n_0} = 0$ implies

$$G_n = \lambda\cdot\alpha^{n_0}\cdot\alpha^{n-n_0} + \mu\cdot\beta^{n_0}\cdot\beta^{n-n_0} = \lambda\cdot\alpha^{n_0}\cdot\sqrt{\Delta}\cdot R_{n-n_0}$$

$$= -\lambda\cdot\beta^{-n_0}\cdot\sqrt{\Delta}\cdot B^n\cdot R_{n_0-n} .$$

Thus we have

$$G_0 = [\![-\lambda\cdot\beta^{-n_0}\cdot\sqrt{\Delta}]\!]\cdot R_{n_0} , \quad G_1 = [\![-\lambda\cdot\beta^{-n_0}\cdot\sqrt{\Delta}]\!]\cdot B\cdot R_{n_0-1} .$$

Suppose that $\mathfrak{p} \mid (R_n, B\cdot R_{n-1})$ for some prime ideal $\mathfrak{p}$ in $\mathbb{Q}(\sqrt{\Delta})$. Then $\mathfrak{p} \mid (\alpha\cdot R_n - B\cdot R_{n-1}) = (\alpha)^n$, and $\mathfrak{p} \mid (\beta\cdot R_n - B\cdot R_{n-1}) = (\beta)^n$, which contradicts $(A,B) = 1$. Thus $(R_n, B\cdot R_{n-1}) = 1$, and then by $(G_0, G_1) = 1$ we must have

$$|\lambda\cdot\beta^{-n_0}\cdot\sqrt{\Delta}| = 1 .$$

Thus we find that $G_n = 0$ implies

$$n = \frac{2}{\log B}\cdot\log|\lambda\cdot\sqrt{\Delta}| < C_3 .$$

Now we turn to the case $G_n \neq 0$. We have from (4.7)

$$\left| \left(\frac{-\lambda}{\mu}\right)\cdot\left(\frac{\alpha}{\beta}\right)^n - 1 \right| \leqslant \frac{v}{|\mu|}\cdot B^{-n/2} . \tag{4.8}$$

We may assume $n \geqslant 2$. Let $-\lambda/\mu = e^{2\pi i\cdot\psi}$, $\alpha/\beta = e^{2\pi i\cdot\varphi}$, with $-\frac{1}{2} < \psi \leqslant \frac{1}{2}$ and $-\frac{1}{2} < \varphi \leqslant \frac{1}{2}$. Let $k \in \mathbb{Z}$ be such that $|\ \psi + n\cdot\varphi + k\ | \leqslant \frac{1}{2}$. Then $|k| \leqslant 1 + \frac{1}{2}\cdot n \leqslant n$. Put

$$\Lambda = 2\pi i\cdot(\ \psi + n\cdot\varphi + k\ ) = \mathrm{Log}\left(\frac{-\lambda}{\mu}\right) + n\cdot\mathrm{Log}\left(\frac{\alpha}{\beta}\right) + 2\cdot k\cdot\mathrm{Log}(-1) .$$

By lemma 2.3 and (4.8) we have an upper bound for $|\Lambda|$:

$$\begin{aligned} |\Lambda| &= 2\pi\cdot|\ \psi + n\cdot\varphi + k\ | \leqslant \tfrac{1}{2}\pi\cdot|e^{2\pi i\cdot(\psi+n\cdot\varphi+k)} - 1| \\ &= \tfrac{1}{2}\pi\cdot\left| \left(\frac{-\lambda}{\mu}\right)\cdot\left(\frac{\alpha}{\beta}\right)^n - 1 \right| \leqslant \tfrac{1}{2}\pi\cdot\frac{v}{|\mu|}\cdot B^{-n/2} . \end{aligned} \tag{4.9}$$

From $G_n \neq 0$ we derive $\Lambda \neq 0$. Then from lemma 2.4 we can derive a lower bound for $|\Lambda|$. Note that $\max(n, 2|k|) \leqslant 2\cdot n$, so that $W = \log(2\cdot n)$. We choose $V_1 = \frac{1}{2}$. The number $z = \alpha/\beta$ satisfies

$$B\cdot z^2 - (A^2-2\cdot B)\cdot z + B = 0 ,$$

hence $h(\alpha/\beta) \leqslant \frac{1}{2}\cdot\log B$. And $z = -\lambda/\mu$ satisfies

$$E\cdot z^2 - (2\cdot E+\Delta\cdot G_0^2)\cdot z + E = 0 ,$$

hence $h(-\lambda/\mu) \leqslant \frac{1}{2}\cdot\log E$. Thus $V_2 = U_2^+$, $V_3 = U_3^+$ satisfy the requirements for Theorem 2.4. We find

$$\begin{aligned} |\Lambda| &> \exp\Big(-C_1\cdot(\ \log(2\cdot n) + \log(2\cdot e\cdot U_3^+)\ )\Big) \\ &= \exp\Big(-C_1\cdot(\ \log n + C_2\ )\Big) . \end{aligned} \tag{4.10}$$

Combining (4.9) and (4.10) we find $n < a + b\cdot\log n$, where

$$a = \frac{2}{\log B}\cdot\left(\log v + \log\frac{\pi}{2\cdot|\mu|} + C_1\cdot C_2\right) ,$$

$$b = 2\cdot C_1/\log B .$$

The result now follows from Lemma 2.1, since

$$b = 2\cdot C_1/\log B = 1.681\times10^{21}\cdot\frac{\max(\pi,\log B)}{\log B}\cdot\max(\pi,\log E)\cdot\log(2\cdot e\cdot U_2^+)$$

which is certainly larger than e^2 . □

Remark. Note that v may depend on n . Thus we can find an upper bound for the solutions $n \in \mathbb{N}_0$ of e.g. $|G_n| \leqslant n^c$ for any constant c .

We now want to reduce the bound found in Theorem 4.3. We do this by studying the diophantine inequality

$$|\ \psi + n\cdot\varphi + k\ | < v_0\cdot B^{-n/2} , \tag{4.11}$$

which follows from (4.9), where $v_0 = v/4\cdot|\mu|$. We have to distinguish between the homogeneous case $\psi = 0$ and the inhomogeneous case $\psi \neq 0$. We apply the methods that have been described in Sections 3.2 and 3.3 respectively. Unlike in other chapters, here we give the results in the form of precisely defined algorithms.

First we study the homogeneous case $\psi = 0$. We then use Algorithm H (see the next page). Let N be an upper bound for n for the solutions of (4.11), for example the bound found in Theorem 4.3.

Input: φ, B, $|\mu|$, v_0, N .

Output: new, reduced bound N^* for n .

(i) (initialization) *Choose* $n_0 \geqslant 2/\log B$ *such that* $B^{n_0/2}/n_0 \geqslant 2\cdot v_0$; $N_0 := N$; *compute the continued fraction*

$$|\varphi| = [\ 0,\ a_1,\ a_2,\ \ldots,\ a_{\ell_0+1},\ \ldots\ ]$$

and the denominators $q_1, \ldots, q_{\ell_0+1}$ *of the convergents of* $|\varphi|$, *with* ℓ_0 *so large that* $q_{\ell_0} \leqslant N_0 < q_{\ell_0+1}$; $i := 0$;

(ii) (compute new bound) $A_i := \max(a_1,\ldots,a_{\ell_i+1})$; *compute the largest integer* N_{i+1} *such that*

$$B^{N_{i+1}/2}/N_{i+1} \leqslant v_0\cdot(A_i+2)\ ,$$

and ℓ_{i+1} *such that* $q_{\ell_{i+1}} \leqslant N_{i+1} < q_{\ell_{i+1}+1}$;

(iii) (terminate loop)

if $n_0 \leqslant N_{i+1} < N_i$ then $i := i+1$, goto (ii) ;
else $N^* := \max(n_0, N_{i+1})$, stop .

Figure 4. ALGORITHM H. (reduces upper bound for (4.11) in the case $\psi = 0$).

LEMMA 4.6. *Algorithm* H *terminates. Inequality* (4.11) *with* $\psi = 0$ *has no solutions with* $N^* < n < N$.

Proof. Termination is obvious, since all N_i are integers. Note that $B^{x/2}/x$ is an increasing function for $x \geqslant 2/\log B$. Hence, if $n \geqslant n_0$,

$$\Big|\ |\varphi| - |k|/n\ \Big| \leqslant v_0\cdot B^{-n/2}/n < 1/2n^2\ .$$

It follows (cf. (3.6)) that $|k|/n$ is a convergent of $|\varphi|$, say $|k|/n = p_m/q_m$. Then $q_m \leqslant n$, and (cf. (3.5)),

$$\Big|\ |\varphi| - p_m/q_m\ \Big| > 1/(a_{m+1}+2)\cdot q_m^2\ .$$

Suppose $n \leqslant N_i$ for some $i \geqslant 0$. Then $m \leqslant \ell_i$. Hence,

$$B^{n/2}/n \leqslant v_0\cdot n^{-2}\cdot\Big|\ |\varphi| - |k|/n\ \Big|^{-1} < v_0\cdot(a_{m+1}+2) \leqslant v_0\cdot(A_m+2)\ .$$

It follows that if $N_{i+1} \geqslant n_0$ then $n \leqslant N_{i+1}$. □

Next we study the inhomogeneous case $\psi \neq 0$. Again, let N be an upper bound for n satisfying (4.11) . We now have the following Algorithm I.

Input: φ, ψ, B, v_0, N .

Output: new, reduced upper bound N^* for all but a finite number of explicitly given n .

(i) (initialization) $N_0 := [N]$; *compute the continued fraction*

$$|\varphi| = [\, 0,\ a_1,\ a_2,\ \dots,\ a_{\ell_0},\ \dots\,]$$

and the convergents p_i/q_i for $i = 1, \dots, \ell_0$, *with* ℓ_0 *so large that* $q_{\ell_0} > 4\cdot N_0$ *and* $\|q_{\ell_0}\cdot\psi\| > 2\cdot N_0/q_{\ell_0}$. *(If such* ℓ_0 *cannot be found within reasonable time, take* ℓ_0 *so large that* $q_{\ell_0} > 4\cdot N_0$ *)* ; $i := 0$;

(ii) (compute new bound)

    <u>if</u> $\|q_{\ell_i}\cdot\psi\| > 2\cdot N_i/q_{\ell_i}$

        <u>then</u> $N_{i+1} := [2\cdot\log(q_{\ell_i}^2\cdot v_0/N_i)/\log B]$;

        <u>else</u> *compute* $K \in \mathbb{Z}$ *with* $|\, K - q_{\ell_i}\cdot\psi\,| \leqslant \frac{1}{2}$; *compute*

            $n_0 \in \mathbb{Z}$, $0 \leqslant n_0 < q_{\ell_i}$, *with* $K = n_0\cdot p_{\ell_i} \equiv 0 \pmod{q_{\ell_i}}$;

            <u>if</u> $n = n_0$ *is a solution of* (4.11), <u>then</u> *print an appropriate message*;

            $N_{i+1} := [2\cdot\log(4\cdot q_{\ell_i}\cdot v_0)/\log B]$;

(iii) (terminate loop)

    <u>if</u> $N_{i+1} < N_i$

        <u>then</u> $i := i + 1$; *compute the minimal* $\ell_i < \ell_{i-1}$ *such that*

            $q_{\ell_i} > 4\cdot N_i$ *and* $\|q_{\ell_i}\cdot\psi\| > 2\cdot N_i/q_{\ell_i}$ *(if such* ℓ_i *does not exist, choose the minimal* ℓ_i *with* $q_{\ell_i} > 4\cdot N_i$ *)*;

            <u>goto</u> (ii) ;

        <u>else</u> $N^* := N_i$; <u>stop</u> .

<u>Figure 5.</u> <u>ALGORITHM I.</u> (reduces upper bound for (4.11) in the case $\psi \neq 0$).

<u>LEMMA 4.7.</u> *Algorithm* I *terminates. Inequality* (4.11) *with* $\psi \neq 0$ *has for* $N^* < n < N$ *only the finitely many solutions found by the algorithm.*

<u>Proof.</u> It is clear that the algorithm terminates. Suppose that $n \leqslant N_i$ for

some $i \geqslant 0$. Then if $\|q_{\ell_i}\cdot\psi\| > 2\cdot N_i/q_{\ell_i}$, we have

$$\|q_{\ell_i}\cdot\psi\| = \|q_{\ell_i}\cdot(\psi+n\cdot\varphi+k) - n\cdot\varphi\cdot q_{\ell_i}\|$$

$$\leqslant q_{\ell_i}\cdot|\psi+n\cdot\varphi+k| + n/q_{\ell_i} \leqslant q_{\ell_i}\cdot v_0\cdot B^{-n/2} + N_i/q_{\ell_i} .$$

It follows that $n \leqslant N_{i+1}$. If $\|q_{\ell_i}\cdot\psi\| \leqslant 2\cdot N_i/q_{\ell_i}$, then

$$|K+n\cdot p_{\ell_i}+k\cdot q_{\ell_i}| \leqslant |K-q_{\ell_i}\cdot\psi| + q_{\ell_i}\cdot|\psi+n\cdot\varphi+k| + n\cdot|p_{\ell_i}-q_{\ell_i}\cdot\varphi|$$

$$\leqslant \tfrac{1}{2} + q_{\ell_i}\cdot v_0\cdot B^{-n/2} + N_i/q_{\ell_i} < \tfrac{3}{4} + q_{\ell_i}\cdot v_0\cdot B^{-n/2} .$$

If $q_{\ell_i}\cdot v_0\cdot B^{-n/2} \leqslant \frac{1}{4}$, then $K + n\cdot p_{\ell_i} + k\cdot q_{\ell_i} = 0$, since it is an integer. By $(p_{\ell_i},q_{\ell_i}) = 1$ it follows that $n \equiv n_0 \pmod{q_{\ell_i}}$. Since $q_{\ell_i} > N_i$, the only possibility is $n = n_0$. If $q_{\ell_i}\cdot v_0\cdot B^{-n/2} > \frac{1}{4}$, then $n \leqslant N_{i+1}$ follows immediately. □

We remark that in practice one almost always finds an ℓ_i such that $\|q_{\ell_i}\cdot\psi\| > 2\cdot N_i/q_{\ell_i}$, if N_i is large enough.

4.4. Upper bounds.

In this section we will derive explicit upper bounds for the solutions of (4.1), both in the hyperbolic and elliptic cases. Our first step is the application of the p-adic theory of linear forms in logarithms, which works the same way in both cases. We use it to find a bound for m_i that is polynomial in $\log n$. Then we combine this with the results of Section 4.3 on the growth of the recurrence sequence, which for the solutions of (4.1) yield a bound for n that is linear in the m_i (Corollaries 4.3 and 4.5).

Assume that $n_0 \geqslant 2$. Let D be the discriminant of $\mathbb{Q}(\sqrt{\Delta})$. Put

$$L = \log\max\left(|e\cdot D|^{1/4},\ |\alpha\cdot\lambda\cdot\sqrt{\Delta}|,\ |\alpha\cdot\mu\cdot\sqrt{\Delta}|,\ |\beta\cdot\lambda\cdot\sqrt{\Delta}|,\ |\beta\cdot\mu\cdot\sqrt{\Delta}| \right) .$$

Let d be the squarefree part of Δ . For $i = 1, \ldots, s$ put

$$\varphi_i = 2 \text{ if } p_i \mid d ,\quad \varphi_i = 1 \text{ otherwise,}$$

$$\rho_i = 2 \quad \text{if} \quad p_i = 2,\ d \equiv 5 \pmod 8 \quad \text{or if} \quad p_i > 2,\ \left(\frac{d}{p_i}\right) = -1 ,$$

$$\rho_i = 1 \quad \text{otherwise},$$

$$C_{4,i} = 10^6 \cdot \left(\frac{2}{\rho_i \cdot \log p_i}\right)^7 \cdot \varphi_i^{-3} \cdot L^4 \cdot p_i^{4\cdot\rho_i+4} \cdot \left(1 + \frac{\varphi_i \cdot L \cdot p_i^{\rho_i} + 2/L}{\log n_0} \right)^3 .$$

<u>LEMMA 4.8.</u> *The solutions of* (4.1) *with* $n \geqslant n_0$ *satisfy*

$$m_i < C_{4,i} \cdot (\log n)^3 \quad \text{for} \quad i = 1, \ldots, s .$$

<u>Proof.</u> Rewrite (4.1), using (4.5), as

$$\left(\frac{\alpha}{\beta}\right)^n - \left(\frac{-\mu}{\lambda}\right) = \frac{w}{\lambda} \cdot \beta^{-n} \cdot \prod_{i=1}^{s} p_i^{m_i} .$$

Then, by (4.6),

$$m_i \leqslant m_i - \mathrm{ord}_{p_i}(\lambda) = \mathrm{ord}_{p_i}\left(\frac{w}{\lambda} \cdot \beta^{-n} \cdot \prod_{i=1}^{s} p_i^{m_i}\right) = \mathrm{ord}_{p_i}\left(\left(\frac{\alpha}{\beta}\right)^n - \left(\frac{-\mu}{\lambda}\right)\right) .$$

Apply Lemma 2.5 (Schinzel's result) with $\xi'' = \alpha$, $\xi' = \beta$, $\chi'' = \mu \cdot \sqrt{\Delta}$, $\chi' = -\lambda \cdot \sqrt{\Delta}$. Then we find, using $\mathrm{ord}_{\mathfrak{p}_i}(\cdot) = \varphi_i \cdot \mathrm{ord}_{p_i}(\cdot)$,

$$m_i < 10^6 \cdot \left(\frac{2}{\rho_i \cdot \log p_i}\right)^7 \cdot \varphi_i^{-3} \cdot L^4 \cdot p_i^{4\cdot\rho_i+4} \cdot \left(\log n + \varphi_i \cdot L \cdot p_i^{\rho_i} + 2/L \right)^3 ,$$

from which the result follows, since $n \geqslant n_0$. □

Put

$$C_4 = \max_i(C_{4,i}) , \quad m = \max_i(m_i) , \quad P = \prod_{i=1}^{s} p_i .$$

In the case $\Delta > 0$, let $n_0 > \max\left(2, \log|\lambda/\mu|/\log|\alpha/\beta|\right)$, and put

$$C_5 = \log P \,/\, \left(\log|\alpha| + \min(0, \log(\gamma/|w|)) \right) ,$$

$$C_6 = \max\left(8 \cdot C_4 \cdot (\log 27 \cdot C_4 \cdot C_5)^3,\ 841 \cdot C_4 \right) .$$

In the case $\Delta < 0$, put

$$C_7 = \max\left\{ C_3 + \frac{4}{\log B} \cdot \log\left(2 \cdot |G_0 \cdot \mu \cdot \sqrt{\Delta}|\right),\right.$$

$$8\cdot\left(\left(C_3+\frac{4\cdot\log|w|}{\log B}\right)^{1/3} + \left(\frac{4\cdot C_4\cdot\log P}{\log B}\right)^{1/3}\cdot\log\left(\frac{108\cdot C_4\cdot\log P}{\log B}\right)\right)^3 \Bigg\} ,$$

$$C_{8,i} = C_{4,i}\cdot(\log C_7)^3 \quad \text{for} \quad i = 1, \ldots, s .$$

Then we have the following result, giving explicit upper bounds for the solutions of (4.1).

THEOREM 4.9. *Let* $n, m_1, \ldots, m_s$ *be a solution of* (4.1).
(i). *If* $\Delta > 0$ *and* $n \geqslant n_0$ *then* $n < C_5\cdot C_6$ *and* $m < C_6$.
(ii). *If* $\Delta < 0$ *then* $n < C_7$ *and* $m_i < C_{8,i}$ *for* $i = 1, \ldots, s$.

Proof. (i). Corollary 4.3 yields $n < C_5\cdot m$. By Lemma 4.8 we now have

$$m < C_4\cdot(\log n)^3 < C_4\cdot(\log C_5\cdot m)^3 .$$

If $C_4\cdot C_5 > (e^2/3)^3$, we apply Lemma 2.1 with $a = 0$, $b = C_4\cdot C_5$, $h = 3$, and we find $m < 8\cdot C_4\cdot(\log 27\cdot C_4\cdot C_5)^3$. If $C_4\cdot C_5 \leqslant (e^2/3)^3$, then

$$n < C_5\cdot m < C_4\cdot C_5\cdot(\log n)^3 \leqslant (e^2/3)^3\cdot(\log n)^3 ,$$

from which we deduce $n < 12564$. Now, $m < C_4\cdot(\log n)^3 < 841\cdot C_4$.
(ii). From Lemma 4.8 and Corollary 4.5 we see that

$$n < C_3 + \frac{4}{\log B}\cdot\log\left(2\cdot|G_0\cdot\mu\cdot\sqrt{\Delta}|\right) ,$$

or

$$n < C_3 + \frac{4\cdot\log|w|}{\log B} + \frac{4\cdot C_4\cdot\log P}{\log B}\cdot(\log n)^3 .$$

The result now follows from Lemma 2.1, since $4\cdot C_4\cdot\log P/\log B > (e^2/3)^3$. □

4.5. A basic lemma.

We introduce some notation, and then give an almost trivial lemma that is at the heart of our reduction methods for both the hyperbolic and the elliptic cases. Let for $i = 1, \ldots, s$

$$e_i = -\mathrm{ord}_{p_i}(\lambda) , \quad f_i = \mathrm{ord}_{p_i}(\log_{p_i}(\tfrac{\alpha}{\beta})) , \quad g_i = f_i - e_i ,$$

$$\vartheta_i = -\log_{p_i}(\tfrac{-\lambda}{\mu})/\log_{p_i}(\tfrac{\alpha}{\beta}) .$$

By Lemma 4.1 the p_i-adic logarithms of α/β and $-\lambda/\mu$ exist. Note that $\log_{p_i}(\alpha/\beta) \neq 0$, since the sequence $\{G_n\}$ is not degenerate. Note that for conjugated ξ, ξ' also $\log_p \xi$ and $\log_p \xi'$ are conjugates, hence $\log_p(\xi/\xi') \in \sqrt{\Delta}\cdot\mathbb{Q}_p$. Hence both numerator and denominator of ϑ_i are in $\sqrt{\Delta}\cdot\mathbb{Q}_{p_i}$, so $\vartheta_i \in \mathbb{Q}_{p_i}$. Hence, if $\vartheta_i \neq 0$, we can write

$$\vartheta_i = \sum_{\ell=k_i}^{\infty} u_{i,\ell}\cdot p_i^{\ell} ,$$

where $k_i = \mathrm{ord}_{p_i}(\vartheta_i)$ and $u_{i,\ell} \in \{ 0, 1, \ldots, p_i-1 \}$ for all ℓ . The following lemma localizes the elements of $\{G_n\}$ with many factors p_i , in terms of the p_i-adic expansion of ϑ_i .

<u>LEMMA 4.10.</u> *Let* $n \in \mathbb{N}_0$. *If* $\mathrm{ord}_{p_i}(G_n) + e_i > 1/(p_i-1)$ *then*

$$\mathrm{ord}_{p_i}(G_n) = g_i + \mathrm{ord}_{p_i}(n-\vartheta_i) .$$

<u>Proof.</u> By Lemma 4.1 we have

$$\mathrm{ord}_{p_i}(G_n) + e_i = \mathrm{ord}_{p_i}\left(\left(\frac{\alpha}{\beta}\right)^n - \left(\frac{-\mu}{\lambda}\right)\right) = \mathrm{ord}_{p_i}\left(\left(\frac{-\lambda}{\mu}\right)\cdot\left(\frac{\alpha}{\beta}\right)^n - 1\right) .$$

With $\xi = (-\lambda/\mu)\cdot(\alpha/\beta)^n - 1$ we have by assumption $\mathrm{ord}_{p_i}(\xi) > 1/(p_i-1)$. Hence $\mathrm{ord}_{p_i}(\xi) = \mathrm{ord}_{p_i}(\log_{p_i}(1+\xi))$, and it follows that

$$\mathrm{ord}_{p_i}(G_n) + e_i = \mathrm{ord}_{p_i}\left(n\cdot\log_{p_i}\left(\frac{\alpha}{\beta}\right) + \log_{p_i}\left(\frac{-\lambda}{\mu}\right) \right)$$

$$= \mathrm{ord}_{p_i}(n-\vartheta_i) + f_i . \qquad \square$$

4.6. Trivial cases.

We have to exclude some trivial cases first. The first trivial case is that of $\mathrm{ord}_{p_i}(\vartheta_i) < 0$. Then the solutions of (4.1) satisfy $m_i \leqslant 1/(p_i-1) - e_i$, or, by Lemma 4.10,

$$m_i = f_i - e_i + \mathrm{ord}_{p_i}(n-\vartheta_i) .$$

Since $n \in \mathbb{Z}$ and $\mathrm{ord}_{p_i}(\vartheta_i) < 0$ we have $\mathrm{ord}_{p_i}(n-\vartheta_i) = \mathrm{ord}_{p_i}(\vartheta_i)$. Hence

$$m_i \leqslant \max\Big(f_i + \mathrm{ord}_{p_i}(\vartheta_i),\ 1/(p_i-1) \Big) - e_i \ .$$

The case where all p_i-adic digits of ϑ_i from a certain point on are all zero is a special case, because the reduction methods of the next sections then do not work. This is so because these reduction methods make use of zero-dimensional p-adic diophantine approximation, as explained in Section 3.9, applied to the p-adic linear form

$$\log_p\left(\frac{\lambda}{\mu}\right) + n\cdot\log_p\left(\frac{\alpha}{\beta}\right)$$

for $p = p_1, \ldots, p_s$. This means that we must study the p-adic number

$$\vartheta = -\log_p\left(\frac{\lambda}{\mu}\right) / \log_p\left(\frac{\alpha}{\beta}\right) \ .$$

If it happens that this number ϑ is zero, or that all digits in the p-adic expansion of ϑ are zero from a certain point on, then obviously the reduction process of Section 3.9 breaks down, since it is based on the assumption that the p-adic expansion of ϑ contains sufficiently many non-zero digits.

This case can be dealt with as follows. Note that $\vartheta_i = r$ holds for all $i = 1, \ldots, s$ with the same r . Thus, by Lemma 4.10,

$$m_i \leqslant \max\Big(g_i + \mathrm{ord}_{p_i}(n-r),\ 1 - e_i \Big) \leqslant g_i + 1 + \mathrm{ord}_{p_i}(n-r) \ . \quad (4.12)$$

Then we have, if $\Delta > 0$, by Corollary 4.3,

$$n\cdot\log|\alpha| < \sum_{i=1}^{s} (g_i+1)\cdot\log p_i - \log(\gamma/|w|) + \log|n-r| \ ,$$

from which a good upper bound for n can be derived (no application of the Gelfond-Baker theory is involved, so the constants are relatively small). And if $\Delta < 0$, the proof of Lemma 4.11 below yields $\vartheta_i = 0$, whence, by (4.12),

$$|G_n| = |w|\cdot\prod_{i=1}^{s} p_i^{m_i} \leqslant v_0\cdot n$$

for some constant v_0 . Only minor changes in the results and algorithms of Section 4.3 suffice to deal with this inequality instead of (4.7).

There is however an elementary way of treating this case, using congruences only, that is guaranteed to work. We define the following special 'symmetric recurrences'. For α, β as defined in Section 4.2, let d be the squarefree part of Δ, and put

$$R_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \quad S_n = \alpha^n + \beta^n,$$

for $d = -1$ also

$$T_n^{\pm} = (1 \pm \sqrt{(-1)})\cdot\alpha^n + (1 \mp \sqrt{(-1)})\cdot\beta^n,$$

and for $d = -3$ also (with $\omega = \rho$ or $\bar{\rho}$ for $\rho = \frac{1}{2}\cdot(1+\sqrt{(-3)})$)

$$U_n(\omega) = (1 + \omega)\cdot\alpha^n + (1 + \bar{\omega})\cdot\beta^n,$$

$$V_n(\omega) = \omega\cdot\alpha^n + \bar{\omega}\cdot\beta^n,$$

for all $n \in \mathbb{Z}$. Note that

$$T_n^{+}\cdot T_n^{-} = 2\cdot S_{2n}, \quad U_n(\omega)\cdot U_n(\bar{\omega})\cdot R_n = 3\cdot R_{3n}, \quad V_n(\omega)\cdot V_n(\bar{\omega})\cdot S_n = S_{3n}.$$

We have the following lemma. We assume that $\mathrm{ord}_p(\vartheta) \geqslant 0$.

<u>LEMMA 4.11.</u> *If* ϑ *has only finitely many nonzero* p-*adic digits, then there exist an* $r \in \mathbb{N}_0$ *and a* $\kappa \in \mathbb{Q}$ *such that* $G_n = \kappa\cdot R_{n-r}$*, or* $G_n = \kappa\cdot S_{n-r}$*, or (if* $d = -1$*)* $G_n = \kappa\cdot T_n^{\pm}$*, or (if* $d = -3$*)* $G_n = \kappa\cdot U_n(\omega)$ *or* $\kappa\cdot V_n(\omega)$*, where* $\omega = \rho$ *or* $\bar{\rho}$*. Further,* $r = 0$ *if* $\Delta < 0$.

<u>Proof.</u> By $\mathrm{ord}_p(\vartheta) \geqslant 0$ we have $\vartheta = r$ for some $r \in \mathbb{N}_0$. From the definition of ϑ we infer

$$\log_p \left(\frac{\alpha}{\beta}\right)^r \cdot \left(\frac{-\lambda}{\mu}\right) = 0,$$

hence $\eta = (\beta/\alpha)^r\cdot(\mu/\lambda)$ is a root of unity. It follows that we can write

$$G_n = \lambda\cdot\alpha^r\cdot(\alpha^{n-r} + \eta\cdot\beta^{n-r}).$$

First let $B = \pm 1$. Then $\Delta > 0$ and

$$G_0 = \lambda\cdot\alpha^r\cdot(\alpha^{-r} \pm \beta^{-r}) = \pm\lambda\cdot\alpha^r\cdot(\alpha^r \pm \beta^r),$$

$$G_1 = \lambda\cdot\alpha^r\cdot(\alpha^{1-r} \pm \beta^{1-r}) = \pm\lambda\cdot\alpha^r\cdot(\alpha^{r-1} \pm \beta^{r-1}).$$

Note that

$$(\alpha^{r-1} + \beta^{r-1},\ \alpha^r + \beta^r) = (2,\ \alpha + \beta) = (1) \quad \text{or} \quad (2) ,$$

$$(\alpha^{r-1} - \beta^{r-1},\ \alpha^r - \beta^r) = (\alpha - \beta) .$$

By $(G_0,G_1) = 1$ it follows that $\pm\lambda\cdot\alpha^r = 1, \frac{1}{2}$ or $1/(\alpha-\beta)$, respectively, and the assertion follows.
Next suppose $|B| \geqslant 2$. Then

$$G_0\cdot B\cdot(\eta\cdot\alpha^{r-1} + \beta^{r-1}) = G_1\cdot(\eta\cdot\alpha^r \pm \beta^r) .$$

Since $(B,G_1) = 1$, we have $\alpha\cdot\beta \mid \eta\cdot\alpha^r \pm \beta^r$. By $(A,B) = 1$ we have $(\alpha,\beta) = (1)$, and from $\alpha \mid \beta^r$ it then follows that $\vartheta = r = 0$. So $G_0 = \lambda\cdot(1+\eta) \in \mathbb{Z}$. The result now follows easily, since for η the only possibilities are ± 1 for all d , and moreover $\pm\sqrt{(-1)}$ if $d = -1$, and $\pm\rho,\ \pm\bar{\rho}$ if $d = -3$. □

In the cases of Lemma 4.11 we can treat (4.1) as follows. Lemma 4.10 shows that the smallest index $n = g(m\cdot p^\ell) > 0$ such that $m\cdot p^\ell \mid G_n$ grows exponentially with ℓ . Also, G_n grows exponentially with n , as follows from Lemma 4.2 and Theorem 4.4. Hence $G_{g(m\cdot p^\ell)}$ grows doubly exponentially with ℓ . It follows that $a = w\cdot p_1^{m_1}\cdot\ldots\cdot p_s^{m_s}$ cannot keep up with $G_{g(a)}$ as the m_i tend to infinity. It follows that if $p_1^{m_1}\cdot\ldots\cdot p_s^{m_s}$ is large enough, there exists a prime q such that $q \mid G_{g(a)}$ but $q \nmid a$. Now the sequences $\{R_n\}$, $\{S_n\}$ have special divisibility properties, such as

$$R_n \mid R_m \quad \text{if and only if} \quad n \mid m ,$$

$$S_n \mid S_{kn} \quad \text{for odd } k ,$$

$$\mathrm{ord}_2(S_n) \leqslant \mathrm{ord}_2(S_3) \quad \text{for all} \quad n \geqslant 1 .$$

Making use of this kind of properties it can be proved that $q \mid G_n$ whenever $a \mid G_n$. This gives an upper bound for the solutions of (4.1), since for those solutions $a \mid G_n$ but $q \nmid G_n$. We give two examples.

Example. Let $A = 16,\ B = 1,\ G_0 = 1,\ G_1 = 8,\ w = 1,\ p_1 = 2,\ p_2 = 11$. Then $\alpha = 8 + 3\cdot\sqrt{7},\ \beta = 8 - 3\cdot\sqrt{7},\ \lambda = \mu = \frac{1}{2}$, so λ/μ is a root of unity. Hence $\vartheta_1 = \vartheta_2 = 0$. Note that we have a sequence of type S_n here. We have

n	-3	-2	-1	0	1	2	3
G_n	2024	127	8	1	8	127	2024
G_n (mod 16)	8	-1	8	1	8	-1	8
G_n (mod 11)	0	6	8	1	8	6	0
G_n (mod 11^2)	88	6	8	1	8	6	88
G_n (mod 23)	0	12	8	1	8	12	0

It follows by this table that $\mathrm{ord}_2(G_n) = 0$ or 3, according to n even or odd, and $\mathrm{ord}_{11}(G_n) > 0$ if and only if $n \equiv 3 \pmod 6$. This can also be derived from Lemma 4.10, which yields: if $\mathrm{ord}_2(G_n) \geqslant 1$ (which happens exactly for odd n), then $\mathrm{ord}_2(G_n) = 3 + \mathrm{ord}_2(n) = 3$. Further, if $\mathrm{ord}_{11}(G_n) \geqslant 1$ (which happens exactly when $n \equiv 3 \pmod 6$), then $\mathrm{ord}_{11}(G_n) = 1 + \mathrm{ord}_{11}(n)$ (e.g. $\mathrm{ord}_{11}(G_{33}) = 2$, but $\mathrm{ord}_{11}(G_{11}) = 0$).

Now, $G_3 \mid G_{3k}$ holds for all odd k. Note that G_3 has exactly 3 factors 2, and 1 factor 11. But it is larger than $2^3\cdot 11 = 88$. Hence there is a prime q, distinct from 2 and 11, such that $q \mid G_n$ whenever $11 \mid G_n$. Thus $G_n = 2^{m_1}\cdot 11^{m_2}$ has no solutions with $m_2 \neq 0$, so that there remain only three solutions: $n = -1, 0, 1$. Note that it is not necessary to know the value of q explicitly. In this case it is 23, and indeed it is easy to show directly that $23 \mid G_n$ if and only if $n \equiv 3 \pmod 6$.

<u>Example.</u> Let $A = 5$, $B = 13$, $G_0 = G_1 = 1$. Then $\Delta = -27$, $\alpha = 1 + 3\cdot\rho$, $\lambda = (1+\rho)/3$. Then $\lambda/\bar{\lambda} = \rho$ is a root of unity, thus $\vartheta = 0$. We will solve $G_n = \pm 2^m$. The sequence $G_n = \lambda\cdot\alpha^n + \bar{\lambda}\cdot\bar{\alpha}^n$ is related to the sequence $H_n = \bar{\lambda}\cdot\alpha^n + \lambda\cdot\bar{\alpha}^n$ and to $R_n = (\alpha^n - \bar{\alpha}^n)/(\alpha - \bar{\alpha})$ by $G_n\cdot H_n\cdot R_n = R_{3n}/3$. Since R_n has nice divisibility properties, we have useful information on the prime divisors of G_n and H_n. We find:

n	0	1	2	3	4	5	6	7	8
G_n	1	1	-8	-53	-161	-116	1513	9073	25696
H_n	1	4	7	-17	-176	-659	-1007	3532	30751
R_n	0	1	5	12	-5	-181	-840	-1847	1685

Now, $G_n \equiv 0 \pmod{16}$ if and only if $n \equiv 8 \pmod{12}$ (Lemma 4.10 yields: if $\mathrm{ord}_2(G_n) \geqslant 2$ (which happens exactly when $n \equiv 2 \pmod 3$), then $\mathrm{ord}_2(G_n) = 2 + \mathrm{ord}_2(n)$), $H_n \equiv 0 \pmod{16}$ if and only if $n \equiv 4 \pmod{12}$, and $R_n \equiv 0 \pmod{16}$ if and only if $n \equiv 0 \pmod{12}$. Note that

$G_4 \cdot H_4 \cdot R_4 = R_{12}/3 = -2^4 \cdot 5 \cdot 7 \cdot 11 \cdot 23$. Considering the sequences modulo 5, 7, 11 and 23 we find that $2^4 \cdot 7 \cdot 11 \cdot 23 \mid G_n \cdot H_n$ for all $n \equiv 0 \pmod 4$, and in fact $11 \mid G_n$ whenever $16 \mid G_n$. Thus $G_n = \pm 2^m$ implies $m \leqslant 3$. It follows from Section 4.3 how to solve $|G_n| \leqslant 8$.

We note that a process as described above can always be applied when dealing with a situation as in Lemma 4.11. This is guaranteed by Lemma 4.10.

From now on we thus assume that $\mathrm{ord}_{p_i}(\vartheta_i) \geqslant 0$ for all $i = 1, \ldots, s$, and that infinitely many p_i-adic digits $u_{i,\ell}$ of ϑ_i are nonzero.

4.7. The reduction algorithm in the hyperbolic case.

First we give the reduction algorithm (Algorithm P, see the next page) for the case $\Delta > 0$. It is based on Lemma 4.10 and Corollary 4.3 only. Let N be an upper bound for n for the solutions $n, m_1, \ldots, m_s$ of (4.1). For example, $N = C_5 \cdot C_6$ as in Theorem 4.9.

THEOREM 4.12. *With all the above assumptions, Algorithm* P *terminates. Equation* (4.1) *with* $\Delta > 0$ *has no solutions with* $N^* \leqslant n < N$, $m_i > M_i$ *for* $i = 1, \ldots, s$.

Proof. Since the p_i-adic expansion of ϑ_i is assumed to be infinite, there exist r_i with the required properties. It is clear that $s_{i,1} \leqslant r_i < s_{i,0}$, and that $N_j \leqslant N_{j-1}$. So $s_{i,j} \leqslant s_{i,j-1}$ holds for all $j \geqslant 1$. Since $s_{i,j} \geqslant 0$, there is a j such that $N_j \leqslant n_0$ or $s_{i,j} = s_{i,j-1}$ for all $i = 1, \ldots, s$. In the latter case, K_j remains .false. ; in both cases the algorithm terminates. We prove by induction on j that $m_i \leqslant g_i + s_{i,j}$ for $i = 1, \ldots, s$, and $n < N_j$ hold for all j . For $j = 0$, it is clear that $n < N_0$. Suppose $n < N_{j-1}$ for some $j \geqslant 1$. Suppose there exists an i such that $m_i > g_i + s_{i,j}$. From Lemma 4.10 we have

$$\mathrm{ord}_{p_i}(n-\vartheta_i) = m_i - g_i \geqslant s_{i,j} + 1 ,$$

hence, by $u_{i,s_{i,j}} \neq 0$,

$$n \geqslant \sum_{\ell=0}^{s_{i,j}} u_{i,\ell} \cdot p^{\ell} \geqslant p^{s_{i,j}} \geqslant N_{j-1} ,$$

Input: $\alpha, \beta, \lambda, \mu, w, p_1, \ldots, p_s, N$.

Output: new, reduced upper bounds M_i for m_i for $i = 1, \ldots, s$, and N^* for n .

(i) (initialization) *Choose an* $n_0 \geqslant 0$ *such that*

$$n_0 > \log|\mu/\lambda|/\log|\alpha/\beta| \;; \quad \gamma := |\lambda| - |\mu|\cdot|\alpha/\beta|^{-n_0} \;;$$

$$\left.\begin{array}{l} g_i := \mathrm{ord}_{p_i}(\lambda) + \mathrm{ord}_{p_i}(\log_{p_i}(\alpha/\beta)) \\[2ex] h_i := \mathrm{ord}_{p_i}(\lambda) + \begin{cases} 3/2 & \text{if } p_i = 2 \\ 1 & \text{if } p_i = 3 \\ 1/2 & \text{if } p_i \geqslant 5 \end{cases} \end{array}\right\} \textit{ for } i = 1, \ldots, s \;;$$

$$g := \gamma \,/\, |w|\cdot\prod_{i=1}^{s} p_i^{g_i} \;; \quad N_0 := N \;;$$

(ii) (computation of the ϑ_i's) *Compute for* $i = 1, \ldots, s$ *the first* r_i p_i*-adic digits* $u_{i,\ell}$ *of*

$$\vartheta_i = -\log_{p_i}\left(\frac{-\lambda}{\mu}\right)/\log_{p_i}\left(\frac{\alpha}{\beta}\right) = \sum_{\ell=0}^{\infty} u_{i,\ell}\cdot p_i^{\ell} \,,$$

where r_i *is so large that* $p_i^{r_i} \geqslant N_0$ *and* $u_{i,r_i} \neq 0$;

(iii) (further initialization, start outer loop) $s_{i,0} := r_i + 1$ *for* $i = 1, \ldots, s$; $j := 1$;

(iv) (start inner loop) $i := 1$; $K_j :=$ <u>.false.</u> ;

(v) (computation of the new bounds for m_i , terminate inner loop)

$s_{i,j} := \min \{ s \in \mathbb{N}_0 \mid p_i^s \geqslant N_{j-1}$ and $u_{i,s} \neq 0 \}$;

<u>if</u> $s_{i,j} < s_{i,j-1}$

 <u>then</u> $K_j :=$ <u>.true.</u> ;

<u>if</u> $i < s$

 <u>then</u> $i := i + 1$; <u>goto</u> (v) ;

(vi) (computation of the new bound for n , terminate outer loop)

$$N_j := \min\left(N_{j-1},\ \left(\sum_{i=1}^{s} s_{i,j}\cdot\log p_i - \log g\right)/\log|\alpha| \right) ;$$

<u>if</u> $N_j \geqslant n_0$ <u>and</u> K_j

 <u>then</u> $j := j + 1$; <u>goto</u> (iv) ;

 <u>else</u> $N^* := \max (N_j, n_0)$;

 $M_i := \max (h_i, g_i + s_{i,j})$ *for* $i = 1, \ldots, s$; <u>stop</u>.

<u>Figure 6.</u> <u>ALGORITHM P.</u> (reduces given upper bounds for (4.1) if $\Delta > 0$).

which contradicts our assumption. Thus, $m_i \leqslant g_i + s_{i,j}$ for $i = 1, \ldots, s$. Then from Corollary 4.3 it follows that

$$n < \left(\sum_{i=1}^{s} (g_i+s_{i,j})\cdot\log p_i - \log(\gamma/|w|) \right)/\log|\alpha| ,$$

hence $n < N_j$. □

Remark 1. In general, one expects that $p_i^{s_{i,j}}$ will not be much larger than N_j, i.e. not too many consecutive p_i-adic digits of ϑ_i will be zero. Then N_j is about as large as $\log N_{j-1}$. In practice, the algorithm will often terminate in three or four steps, near to the largest solution. The computation time is polynomial in s, the bottleneck of the algorithm is the computation of the p_i-adic logarithms.

Remark 2. Pethö [1985] gives for $s = 1$ a different reduction algorithm. For a prime p_i he computes the function $g(u)$, defined for $u \in \mathbb{N}$ as the smallest index $n \geqslant 0$ such that $G_n \neq 0$ and $p_i^u \mid G_n$. Note that if the p_i-adic limit $\lim_{u\to\infty} g(u)$ exists, then by Lemma 4.10 it is equal to ϑ_i.

Remark 3. If $B = \pm 1$ (hence $\Delta > 0$), we can extend the sequence $\{G_n\}_{n=0}^{\infty}$ to negative indices by the recursion formula

$$G_{n-1} = A\cdot B\cdot G_n - B\cdot G_{n+1} \quad \text{for} \quad n = 0, -1, -2, \ldots$$

(cf. (4.3)). Then (4.5) is true for $n < 0$ also. We can solve equation (4.1) with $n \in \mathbb{Z}$ not necessarily nonnegative, by applying Algorithm P twice: once for $\{G_n\}_{n=0}^{\infty}$, and once for the sequence $\{G'_n\}_{n=0}^{\infty}$, defined by $G'_n = G_{-n}$. Note that $G'_n = B^n\cdot\left(\mu\cdot\alpha^n+\lambda\cdot\beta^n\right)$, and

$$\vartheta'_i = -\frac{\log_{p_i}(-\mu/\lambda)}{\log_{p_i}(\alpha/\beta)} = +\frac{\log_{p_i}(-\lambda/\mu)}{\log_{p_i}(\alpha/\beta)} = -\vartheta_i \quad \text{for} \quad i = 1, \ldots, s .$$

Now, instead of applying Algorithm P twice, we can modify it, so that it works for all $n \in \mathbb{Z}$, as follows. Lemmas 4.8 and 4.10 remain correct if we replace n by $|n|$. In Theorem 4.9 the lower bound for n_0 must be replaced by

$$n_0 > \max\left(2,\ |\log|\mu/\lambda||/\log|\alpha/\beta|,\ |\log|\lambda/\mu||/\log|\alpha/\beta| \right) ,$$

and γ has to be replaced by

$$\gamma = \min\left(|\lambda| - |\mu|\cdot|\alpha/\beta|^{-n_0},\ |\mu| - |\lambda|\cdot|\alpha/\beta|^{-n_0} \right) .$$

Similar modifications should be made in step (i) of Algorithm P. Further, in step (ii), r_i should be chosen so large that

$$\underline{\text{if}}\ \ p_i \neq 2\ \ \underline{\text{then}}\ \ p_i^{r_i} \geqslant N_0 \ \text{ and }\ u_{i,r_i} \neq 0\ ,\ \ u_{i,r_i} \neq p-1\ ;$$

$$\underline{\text{else}}\ \ p_i^{r_i-1} \geqslant N_0 \ \text{ and }\ u_{i,r_i} \neq u_{i,r_i-1}\ ;$$

and similar modifications have to be made in step (v) for $s_{i,j}$. With these changes, Theorem 4.12 remains true with n replaced by $|n|$.

We conclude this section with an example.

Example. Let $A = 6,\ B = 1,\ G_0 = 1,\ G_1 = 4,\ w = 1,\ p_1 = 2,\ p_2 = 11$. Then $\alpha = 3 + 2\cdot\sqrt{2},\ \beta = 3 - 2\cdot\sqrt{2},\ \lambda = (\ 1 + 2\cdot\sqrt{2}\ )/4\cdot\sqrt{2},\ \mu = (\ -1 + 2\cdot\sqrt{2}\ )/4\cdot\sqrt{2}$, and $\Delta = 32$. With $n_0 = e^{60} = 1.142\times10^{26}$ we find $C_4 < 2.49\times10^{20}$. With the modifications of Remark 3 above we have $\gamma > 0.323$, $C_5 < 1.76$, $C_6 < 2.62\times10^{26}$, $C_5\cdot C_6 < 4.62\times10^{26}$. Hence all solutions of $G_n = 2^{m_1}\cdot 11^{m_2}$ satisfy $|n| < 4.62\times10^{26}$, $\max(m_1,m_2) < 2.62\times10^{26}$. We perform the reduction Algorithm P step by step. (We write the p-adic number $\sum_{\ell=0}^{\infty} u_\ell\cdot p^\ell$ as $0.u_0u_1u_2\ldots.$, and if $p > 10$ we denote the digits larger than 9 by the symbols A, B, C, ...).

(i) $n_0 = 2,\ \gamma > 0.303,\ g_1 = 0,\ g_1 = 1,\ g > 0.0275,$

$h_1 = -1,\ h_2 = \frac{1}{2},\ N_0 = 4.62\times10^{26}$.

(ii) $\vartheta_1 =$ 0.10111 10111 01000 11100 10100 01001 10001 10010
00001 11101 01000 10000 01001 10011 10101 01101
11100 01011 00001 11010 00011 01001 01010 00101
10001 01011 00000 11001 01011 11101 10100 01011
001.... ,

$\vartheta_2 =$ 0.A9359 05530 7330A 1A223 96230 3A006 A3366 83368
8270.... ,

so $r_1 = 90$ (since $u_{1,89} = 1,\ u_{1,90} = 0,\ 2^{89} > N_0$),

$r_2 = 29$ (since $u_{2,29}) = 6,\ 11^{29} > N_0$).

(iii) $s_{1,0} = 91,\ s_{2,0} = 30$;

(v)-(vi) $s_{1,1} = 90,\ s_{2,1} = 29,\ K_1 =$ .true., $N_1 < 76.9$;

(v)-(vi) $s_{1,2} = 10$, $s_{2,2} = 2$, $K_2 =$ .true., $N_2 < 8.7$;

(v)-(vi) $s_{1,3} = 6$, $s_{2,3} = 1$, $K_3 =$ .true., $N_3 < 5.8$;

(v)-(vi) $s_{1,4} = 6$, $s_{2,4} = 1$, $K_4 =$ .false., $N_4 < 5.8$.

Hence $|n| \leqslant 5$, $m_1 \leqslant 6$, $m_2 \leqslant 2$. We have

n	-5	-4	-3	-2	-1	0	1	2	3	4	5
G_n	2174	373	64	11	2	1	4	23	134	781	4552

So there are 5 solutions: with $n = -3, -2, -1, 0, 1$.

4.8. The reduction algorithm in the elliptic case.

We now present an algorithm to reduce upper bounds for the solutions of (4.1) in the case $\Delta < 0$. The idea is to apply alternatingly Algorithms P and one of H and I. Let N be an upper bound for n , for example $n = C_7$ as in Theorem 4.9.

Input: $\alpha, \beta, \lambda, \mu, w, p_1, \ldots, p_s, N$.
Output: new, reduced upper bounds N^* for n , and M_i for m_i for $i = 1, \ldots, s$.

(i) (initialization) $N_0 := [N]$; $j := 1$;

$$\left.\begin{aligned} g_i &:= \operatorname{ord}_{p_i}(\lambda) + \operatorname{ord}_{p_i}(\log_{p_i}(\alpha/\beta)) \\ h_i &:= \operatorname{ord}_{p_i}(\lambda) + \begin{cases} 3/2 & \text{if } p_i = 2 \\ 1 & \text{if } p_i = 3 \\ 1/2 & \text{if } p_i \geqslant 5 \end{cases} \end{aligned}\right\} \textit{ for } i = 1, \ldots, s \text{ ;}$$

(ii) (computation of the ϑ_i's, φ, ψ) *Compute for* $i = 1, \ldots, s$ *the first* r_i p_i*-adic digits* $u_{i,\ell}$ *of*

$$\vartheta_i = -\log_{p_i}\left(\frac{-\lambda}{\mu}\right)/\log_{p_i}\left(\frac{\alpha}{\beta}\right) = \sum_{\ell=0}^{\infty} u_{i,\ell} \cdot p_i^{\ell} ,$$

where r_i *is so large that* $p_i^{r_i} \geqslant N_0$ *and* $u_{i,r_i} \neq 0$; *compute* $\psi = \operatorname{Log}(-\lambda/\mu)/2\pi i$, *and the continued fraction*

$$|\varphi| = \left|\frac{1}{2\pi i} \cdot \operatorname{Log}(\alpha/\beta)\right| = [\ 0,\ a_1,\ \ldots,\ a_{\ell_0},\ \ldots\ ]$$

with the convergents p_i/q_i *for* $i = 1, \ldots, \ell_0$, *where* ℓ_0 *is so large that* $q_{\ell_0-1} \leqslant N_0 < q_{\ell_0}$ *if* $\psi = 0$; $q_{\ell_0} > 4 \cdot N_0$ *and* $\|q_{\ell_0}\| > 2 \cdot N_0/q_{\ell_0}$ *if* $\psi \neq 0$ *and such* ℓ_0 *can be found in a reasonable amount of time,* $q_{\ell_0} > 4 \cdot N_0$ *otherwise;*

(iii) (one step of Algorithm P) *For* $i = 1, \ldots, s$ *put*

$M_{i,j} := \max\left(h_i,\ g_i + \min\{ s \in \mathbb{N}_0 \mid p_i^s \geqslant N_{j-1} \text{ and } u_{i,s} \neq 0 \} \right)$;

(iv) (one step of Algorithm H or I)

<u>if</u> $\psi = 0$

<u>then</u> $A := \max(a_1, \ldots, a_{\ell_j - 1})$; $v := |w| \cdot \prod_{i=1}^{s} p_i^{M_{i,j}}$;

choose $n_0 \geqslant 2/\log B$ *such that* $B^{n_0/2}/n_0 \geqslant v/2 \cdot |\mu|$;

compute the largest integer N_j *such that*

$B^{N_j/2}/N_j \leqslant (A+2) \cdot v/4 \cdot |\mu|$;

$N_j := \max(n_0, N_j)$;

<u>if</u> $N_j < N_{j-1}$ <u>then</u> *compute* ℓ_j *with* $q_{\ell_j - 1} \leqslant N_j < q_{\ell_j}$;

$j := j + 1$; <u>goto</u> (iii) ;

<u>else</u> <u>if</u> $\|q_{\ell_{j-1}} \cdot \psi\| > 2 \cdot N_{j-1}/q_{\ell_{j-1}}$

<u>then</u> $N_j := [2 \cdot \log\left(q_{\ell_{j-1}}^2 \cdot v/4 \cdot |\mu| \cdot N_{j-1}\right)/\log B]$;

<u>else</u> *compute* $K \in \mathbb{Z}$ *with* $|K - q_{\ell_{j-1}} \cdot \psi| \leqslant \frac{1}{2}$;

compute $n_0 \in \mathbb{Z}$, $0 \leqslant n_0 < q_{\ell_{j-1}}$, *with*

$K + n_0 \cdot p_{\ell_{j-1}} \equiv 0 \pmod{q_{\ell_{j-1}}}$;

<u>if</u> $n = n_0$ *is a solution of* (4.1)

<u>then</u> *print an appropriate message;*

$N_j := [2 \cdot \log\left(q_{\ell_{j-1}} \cdot v/|\mu|\right)/\log B]$;

<u>if</u> $N_j < N_{j-1}$

<u>then</u> *compute the minimal* $\ell_j < \ell_{j-1}$ *such that*

$q_{\ell_j} > 4 \cdot N_j$ *and* $\|q_{\ell_j} \cdot \psi\| > 2 \cdot N_j/q_{\ell_j}$ (*if such* ℓ_j *does not exist, choose the minimal* ℓ_j *such that* $q_{\ell_j} > 4 \cdot N_j$) ; $j := j + 1$; <u>goto</u> (iii) ;

(v) (termination) $N^* := N_{j-1}$; $M_i := M_{i,j}$ *for* $i = 1, \ldots, s$; <u>stop</u>.

<u>Figure 7.</u> <u>ALGORITM C.</u> (reduces upper bounds for (4.1) in the case $\Delta < 0$).

The following theorem now follows at once from the proofs of Lemmas 4.6, 4.7 and Theorem 4.12.

THEOREM 4.13. *Algorithm* C *terminates. Equation* (4.1) *with* $\Delta < 0$ *has no solutions with* $N^* < n < N$ *and* $m_i > M_i$ *for* $i = 1, \ldots, s$, *apart from those spotted by the algoritm.*

We conclude this section with an example.

Example. Let $A = 1$, $B = 2$, $G_0 = 2$, $G_1 = 3$, then $\Delta = -7$, $\alpha = (1 + \sqrt{-7})/2$ and $\lambda = (2 + \sqrt{-7})/\sqrt{-7}$. Let $w = \pm 1$, $p_1 = 3$, $p_2 = 7$. We have with $n_0 = 2$ the following results: $C_4 < 6.40\times10^{16}$, $C_3 < 9.14\times10^{29}$, $C_7 < 7.42\times10^{30}$, $\max(C_{8,1},C_{8,2}) < 2.30\times10^{22}$. Further, $g_1 = 1$, $g_2 = 0$, $h_1 = 1$, $h_2 = 0$. By Theorem 4.9 we may choose $N_0 = 7.42\times10^{30}$. We have

$$\begin{aligned}\varphi = \| \pi - \arctan(\sqrt{7/3}) \| / 2\pi \\ = [\, 0,\ & 2,\ 1,\ 1,\ 2,\ 16,\ 6,\ 1,\ 2,\ 2,\ 13, \\ & 1,\ 1,\ 3,\ 1,\ 1,\ 2,\ 1,\ 2,\ 1,\ 1, \\ & 1,\ 1,\ 1,\ 9,\ 2,\ 1,\ 2,\ 1,\ 7,\ 1, \\ & 6, 269,\ 4,\ 3,\ 1,\ 1,\ 50,\ 2,\ 1,\ 6, \\ & 1,\ 1,\ 2,\ 1,\ 1,\ 7,\ 1,\ 61,\ 1,\ 12, \\ & 3,\ 7,\ 4,\ 7,\ 3, 121,\ 1,\ 21,\ 2,\ 1,\ 7,\ \ldots\ ]\ ,\end{aligned}$$

$$\begin{aligned}\psi &= \| \pi - \arctan(4\cdot\sqrt{7/3}) \| / 2\pi \\ &= 0.29396\ 28336\ 99645\ 40267\ 89566\ 60520\ 01908\ 06203\ldots\ ,\end{aligned}$$

$$\begin{aligned}\vartheta_1 = 0.&20010\ 12210\ 00011\ 02102\ 00211\ 00222\ 02220\ 12021 \\ &10020\ 20202\ 21102\ 00121\ 01000\ 01002\ 11100\ 20122 \\ &11111\ 22202\ 21021\ 02212\ 2200\ldots\ ,\end{aligned}$$

$$\begin{aligned}\vartheta_2 = 0.&32542\ 12042\ 43561\ 34020\ 61561\ 13452\ 10116\ 33152 \\ &25336\ 45044\ 11254\ 55033\ldots\ .\end{aligned}$$

Now we choose $\ell_0 = 61$, since

$$q_{61} = 142\ 51183\ 31142\ 44361\ 19375\ 51238\ 81743 > 4\cdot N_0\ ,$$

and $\|q_{61}\cdot\psi\| = 0.24487\ldots > 2\cdot N_0/q_{61} = 0.104\ldots$. We have $M_{1,1} = 67$, $M_{2,1} = 37$, and we find $N_1 = 637$. Next we choose $\ell_1 = 9$, since $q_9 = 10102 > 4\times637$ and $\|q_9\cdot\psi\| = 0.38745\ldots > 2\times637/10102$. We have $M_{1,2} = 7$, $M_{2,2} = 4$, and we find $N_2 = 74$. Next we choose $\ell_2 = 6$, since $q_6 = 1291 > 4\times74$, and $\|q_6\cdot\psi\| = 0.49398 > 2\times74/1291$. We have $M_{1,3} = 6$,

$M_{2,3} = 3$, and we find $N_3 = 60$. In the next step we find no improvement. Hence $n \leqslant 60$, $m_1 \leqslant 6$, $m_2 \leqslant 3$. It is a matter of straightforward computation to check that there are only the following 6 solutions of $G_n = \pm 3^{m_1} \cdot 7^{m_2}$: $G_1 = 3$, $G_2 = -1$, $G_3 = -7$, $G_5 = 3^2$, $G_7 = 1$, $G_{17} = 3^2 \cdot 7^2$.

4.9. The generalized Ramanujan-Nagell equation.

The most interesting application of the reduction algorithms of the preceding section seems to be the solution of the generalized Ramanujan-Nagell equation (4.2). Let k be a nonzero integer, and let $p_1, \ldots, p_s$ be distinct prime numbers. Then we ask for all nonnegative integers $x, z_1, \ldots, z_s$ with

$$x^2 + k = \prod_{i=1}^{s} p_i^{z_i} .$$

First we note that $z_i = 0$ whenever $-k$ is a quadratic nonresidue $(\text{mod } p_i)$. Thus we assume that this is not the case for all i . Let $p_i \mid k$ for $i = 1, \ldots, t$ and $p_i \nmid k$ for $i = t+1, \ldots, s$. Let $\text{ord}_{p_i}(k)$ be odd for $i = 1, \ldots, r$ and even for $i = r+1, \ldots, t$. Dividing by large enough powers of p_i for $i = 1, \ldots, t$, (4.2) reduces to a finite number of equations

$$D_0 \cdot x_1^2 + k_1 = \prod_{i=r+1}^{s} p_i^{z'_i} \tag{4.13}$$

with $p_i \nmid k_1$ for $i = 1, \ldots, s$, and D_0 composed of $p_1, \ldots, p_r$ only, and squarefree. We distinguish between the 2^{s-r} combinations of z'_i odd or even for $i = r+1, \ldots, s$. Suppose that z'_i is odd for $i = r+1, \ldots, u$ and even for $i = u+1, \ldots, s$. Put

$$y = \prod_{i=r+1}^{u} p_i^{(z'_i-1)/2} \cdot \prod_{i=u+1}^{s} p_i^{z'_i/2} . \tag{4.14}$$

Then, from (4.13),

$$D_0 \cdot x_1^2 - \Big(\prod_{i=r+1}^{u} p_i \Big) \cdot y^2 = -k_1 . \tag{4.15}$$

Put $D = D_0 \cdot \prod_{i=r+1}^{u} p_i$. Then (4.14) and (4.15) lead to

$$\begin{cases} v^2 - D\cdot w^2 = k_2 \\ v = \prod_{i=r+1}^{s} p_i^{m_i} \end{cases}, \tag{4.16}$$

with $v = y\cdot\prod_{i=r+1}^{u} p_i$, $w = x_1$, $k_2 = k_1\cdot\prod_{i=r+1}^{u} p_i$, and also to

$$\begin{cases} v^2 - D\cdot w^2 = k_2 \\ w = \prod_{i=r+1}^{s} p_i^{m_i} \end{cases}, \tag{4.17}$$

with $v = D_0\cdot x_1$, $w = y$, $k_2 = -k_1\cdot D_0$. We proceed with either (4.16) or (4.17), which is the most convenient (e.g. the one with the smaller $|k_2|$).

If $D = 1$, then (4.16) and (4.17) are trivial. So assume $D > 1$. Let ε be the smallest unit in $\mathbb{Z} + \sqrt{D}\cdot\mathbb{Z}$ with $\varepsilon > 1$. It is well known that the solutions v, w of $v^2 - D\cdot w^2 = k_2$ fall apart into a finite number of classes of associated solutions. Let there be T such classes, and choose for $\tau = 1, \ldots, T$ in the τ th class the solution $v_{\tau,0}, w_{\tau,0}$ such that $\gamma_t = v_{\tau,0} + w_{\tau,0}\cdot\sqrt{D} > 1$ is minimal. Then all solutions of $v^2 - D\cdot w^2 = k_2$ are given by $v = \pm v_{\tau,n}$, $w = \pm w_{\tau,n}$, with

$$\begin{cases} v_{\tau,n} = (\ \gamma_\tau\cdot\varepsilon^n + \gamma'_\tau\cdot\varepsilon^{-n}\ )/2 \\ w_{\tau,n} = (\ \gamma_\tau\cdot\varepsilon^n - \gamma'_\tau\cdot\varepsilon^{-n}\ )/2\cdot\sqrt{D} \end{cases} \tag{4.18}$$

for $n \in \mathbb{Z}$, where $\gamma'_t = v_{\tau,0} - w_{\tau,0}\cdot\sqrt{D}$. That is, $\{v_{\tau,n}\}_{n=-\infty}^{\infty}$ and $\{w_{\tau,n}\}_{n=-\infty}^{\infty}$ are linear binary recurrence sequences. Now, (4.16) and (4.17) reduce to T equations of type (4.1). If $k_2 = 1$, then $T = 1$, $\gamma_1 = \varepsilon$, $\gamma'_1 = \varepsilon^{-1}$. If $k_2 \mid 2\cdot D$, $k_2 \neq 1$, then it is easy to prove that $\gamma_\tau^2 = |k_2|\cdot\varepsilon$, $\gamma'^2_t = |k_2|\cdot\varepsilon^{-1}$, so that

$$v_{\tau,n} = \sqrt{|k_2|}\cdot\left(\ (\gamma_\tau/\sqrt{|k_2|})^{2n+1} + (\gamma'_\tau/\sqrt{|k_2|})^{2n+1}\ \right)/2\ ,$$

$$w_{\tau,n} = \sqrt{|k_2|}\cdot\left(\ (\gamma_\tau/\sqrt{|k_2|})^{2n+1} - (\gamma'_\tau/\sqrt{|k_2|})^{2n+1}\ \right)/2\cdot\sqrt{D}\ .$$

In both cases, (4.16) and (4.17) can be solved by elementary means (see Section 4.6, of related interest are Størmer [1897], Mahler [1935], Lehmer [1964], Rumsey and Posner [1964] and Mignotte [1985]). If $k_2 \nmid 2\cdot D$, then we apply the reduction algorithm to one of the equations $v_{\tau,n} = \prod_{i=r+1}^{s} p_i^{m_i}$,

$w_{\tau,n} = \prod_{i=r+1}^{s} p_i^{m_i}$. Note that n is allowed to be negative, since $B = \pm 1$, so we can use the modified algorithm of Remark 3, Section 4.7.

Thus we have a procedure for solving (4.2) completely. It is well known how the unit ε and the minimal solutions $v_{\tau,0}$, $w_{\tau,0}$ for $\tau = 1, \ldots, T$ can be computed by the continued fraction algorithm for $\sqrt{D}$. We conclude this section with an example. It extends the result of Nagell [1948] (also proved by many others) on the original Ramanujan-Nagell equation $x^2 + 7 = 2^z$.

<u>THEOREM 4.14.</u> *The only nonnegative integers* x *such that* $x^2 + 7$ *has no prime divisors larger than* 20 *are the* 16 *in the following table.*

x	$x^2 + 7$	x	$x^2 + 7$	x	$x^2 + 7$
0	7	7	$56 = 2^3 \cdot 7$	31	$968 = 2^3 \cdot 11^2$
1	$8 = 2^3$	9	$88 = 2^3 \cdot 11$	35	$1232 = 2^4 \cdot 7 \cdot 11$
2	11	11	$128 = 2^7$	53	$2816 = 2^8 \cdot 11$
3	$16 = 2^4$	13	$176 = 2^4 \cdot 11$	75	$5632 = 2^9 \cdot 11$
5	$32 = 2^5$	21	$448 = 2^6 \cdot 7$	181	$32768 = 2^{15}$
				273	$74536 = 2^3 \cdot 7 \cdot 11^3$

<u>Proof.</u> Since -7 is a quadratic nonresidue modulo 3, 5, 13, 17 and 19 , we have only the primes 2, 7 and 11 left. Only one factor 7 can occur in $x^2 + 7$, thus we have to solve the two equations

$$x^2 + 7 = 2^{z_1} \cdot 11^{z_2} , \tag{4.19}$$

$$x^2 + 7 = 7 \cdot 2^{z_1} \cdot 11^{z_2} . \tag{4.20}$$

Equation (4.20) can be solved in an elementary way. We distinguish four cases, each leading to an equation of the type

$$y^2 - D \cdot z^2 = c$$

with $c \mid 2 \cdot D$, and either y or z composed of factors 2 and 11 only. We have:

(i) z_1 even, z_2 even, $y = 2^{z_1/2} \cdot 11^{z_2/2}$, $z = x/7$, $c = 1$, $D = 7$;

(ii) z_1 odd, z_2 even, $y = 2^{(z_1+1)/2} \cdot 11^{z_2/2}$, $z = x/7$, $c = 2$, $D = 14$;

(iii) z_1 even, z_2 odd, $y = x$, $z = 2^{z_1/2} \cdot 11^{(z_2-1)/2}$, $c = -7$, $D = 77$;

(iv) z_1 odd, z_2 odd, $y = x$, $z = 2^{(z_1-1)/2} \cdot 11^{(z_2-1)/2}$, $c = -7$, $D = 154$.

In the first example of Section 4.5 we have worked out case (i). We leave the other cases to the reader.

Equation (4.19) can be solved by the reduction algorithm. Again we have four cases, each leading to an equation of the type

$$y^2 - D \cdot z^2 = c$$

with z composed of factors 2 and 11 only. We have

(i) z_1 even, z_2 even, $y = x$, $z = 2^{z_1/2} \cdot 11^{z_2/2}$, $c = -7$, $D = 1$;

(ii) z_1 odd, z_2 even, $y = x$, $z = 2^{(z_1-1)/2} \cdot 11^{z_2/2}$, $c = -7$, $D = 2$;

(iii) z_1 even, z_2 odd, $y = x$, $z = 2^{z_1/2} \cdot 11^{(z_2-1)/2}$, $c = -7$, $D = 11$;

(iv) z_1 odd, z_2 odd, $y = x$, $z = 2^{(z_1-1)/2} \cdot 11^{(z_2-1)/2}$, $c = -7$, $D = 22$.

Case (i) is trivial. The other three cases each lead to one equation of type (4.1). In the example in Section 4.7 we have worked out case (ii). With the following data the reader should be able to perform Algorithm P by hand for the cases (iii) and (iv), thus completing the proof. In these cases $N < 10^{30}$ is a correct upper bound.

Case (iii): $\alpha = 10 + 3\cdot\sqrt{11}$, $\lambda = (2 + \sqrt{11})/2\cdot\sqrt{11}$,

ϑ_1 = 0.10011 01000 00110 10100 00110 10110 01001 11110
11011 10010 00001 10110 10111 10100 00110 01101
01010 10010 11101 11001 10000 10010 01010 11011
00010 00111 01110 00101 01101 01111 10101 11110 10.... ,

ϑ_2 = 0.23075 76425 39004 26090 A92A1 03757 07314 58414 7A238.... .

Case (iv): $\alpha = 197 + 42\cdot\sqrt{22}$, $\lambda = (9 + 2\cdot\sqrt{22})/2\cdot\sqrt{22}$,

ϑ_1 = 0.11101 01101 01110 01010 10111 10001 00100 00011
10000 00110 10101 01100 01101 01111 01101 10101
01011 10100 01100 11101 10011 00011 00010 11110
10101 01100 10011 11111 01001 01110 00000 01110 011.... ,

ϑ_2 = 0.6A001 68184 22921 902A0 724A4 16769 45650 16482 5A6AA.... .

□

Remarks. 1. The computation time for the above proof was less than 2 sec.
2. Let $\Phi(X,Y) = a\cdot X^2 + b\cdot X\cdot Y + c\cdot Y^2$ be a quadratic form with integral coefficients, and $\Delta = b^2 - 4\cdot a\cdot c$ positive or negative. Let k be a nonzero integer, and $p_1, \ldots, p_s$ distinct prime numbers. Then we note that

$$4\cdot a\cdot\Phi(X,Y) = (2\cdot a\cdot X+b\cdot Y)^2 - \Delta\cdot Y^2 ,$$

so that the diophantine equations

$$\Phi(X,k) = \prod_{i=1}^{s} p_i^{z_i} , \quad \Phi(X, \prod_{i=1}^{s} p_i^{z_i}) = k$$

in integers $X \neq 0$ and $z_1, \ldots, z_s \in \mathbb{N}_0$, can both be solved by our method.

4.10. A mixed quadratic-exponential equation.

In this section we give an application of Algorithm C to the following diophantine equation. Let

$$\Phi(X,Y) = a\cdot X^2 + b\cdot X\cdot Y + c\cdot Y^2$$

be a quadratic form with integral coefficients, such that $D = b^2 - 4\cdot a\cdot c$ is negative. Let q, v, w be nonzero integers, and $p_1, \ldots, p_s$ distinct prime numbers. Consider the equation

$$\Phi(X,w\cdot \prod_{i=1}^{s} p_i^{m_i}) = v\cdot q^n \tag{4.21}$$

in integers X , and $n, m_1, \ldots, m_s \in \mathbb{N}_0$.

Let β, $\bar{\beta}$ be the roots of $\Phi(x,1) = 0$. Let h be the class number of $\mathbb{Q}(\sqrt{D})$. There exists a $\pi \in \mathbb{Q}(\sqrt{D})$ such that we have the principal ideal equation $(\pi)\cdot(\bar{\pi}) = (q^h)$. Put $n = n_1 + h\cdot n_2$, with $0 \leqslant n_1 < h$. Then $\Phi(X,Y) = v\cdot q^n$ is equivalent to finitely many ideal equations

$$(a\cdot X-a\cdot\beta\cdot Y)\cdot(a\cdot X-a\cdot\bar{\beta}\cdot Y) = (\sigma)\cdot(\bar{\sigma})\cdot(\pi)^{n_2}\cdot(\bar{\pi})^{n_2} ,$$

with $(\sigma)\bullet(\bar{\sigma}) = (a\cdot v\cdot q^{n_1})$. Hence we have the equations in algebraic numbers

$$\begin{cases} a\cdot X - a\cdot\beta\cdot Y = \gamma\cdot\pi^{n_2} \\ a\cdot X - a\cdot\bar{\beta}\cdot Y = \bar{\gamma}\cdot\bar{\pi}^{n_2} \end{cases} , \quad \begin{cases} a\cdot X - a\cdot\beta\cdot Y = \gamma\cdot\bar{\pi}^{n_2} \\ a\cdot X - a\cdot\bar{\beta}\cdot Y = \bar{\gamma}\cdot\pi^{n_2} \end{cases} ,$$

where γ is composed of σ, units, and common divisors of $a\cdot X - a\cdot\beta\cdot Y$ and $a\cdot X - a\cdot\bar{\beta}\cdot Y$. Note that there are only finitely many choices for γ possible. Thus, (4.21) is equivalent to a finite number of equations

$$a\cdot(\bar{\beta}-\beta)\cdot w\cdot \prod_{i=1}^{s} p_i^{m_i} = \gamma\cdot\pi^{n_2} - \bar{\gamma}\cdot\bar{\pi}^{n_2} ,$$

or, if we put $\lambda = \gamma/a\cdot(\bar{\beta}-\beta)$ and $G_{n_2} = \lambda\cdot\pi^{n_2} + \bar{\lambda}\cdot\bar{\pi}^{n_2}$,

$$G_{n_2} = w\cdot \prod_{i=1}^{s} p_i^{m_i} . \tag{4.22}$$

Here, $\{G_{n_2}\}_{n_2=0}^{\infty}$ is a recurrence sequence with negative discriminant. So (4.22) is of type (4.1), and can thus be solved by the reduction algorithm of Section 4.8.

Before giving an example we remark that (4.21) with $D > 0$ is not solvable with the methods of this chapter. This is due to the fact that in $\mathbb{Q}(\sqrt{D})$ with $D > 0$ there are infinitely many units, hence infinitely many possibilities for γ. Another generalization of equation (4.21) is to replace q^n by $\prod_{i=1}^{t} q_i^{n_i}$. This problem is also not solvable by the method of this chapter, since it does not lead to a binary recurrence sequence if $t \geqslant 2$. These problems can however be dealt with by using multi-dimensional approximation methods, as presented in Chapter 3 and applied in Chapter 7.

We finally present an example.

<u>THEOREM 4.15.</u> *The equation*

$$X^2 - 3^{m_1}\cdot 7^{m_2}\cdot X + 2\cdot\left(3^{m_1}\cdot 7^{m_2}\right)^2 = 11\cdot 2^n$$

in $X \in \mathbb{Z}$, $n, m_1, m_2 \in \mathbb{N}_0$ *has only the following* 24 *solutions:*

n	m_1	m_2	X	n	m_1	m_2	X
1	1	0	−1, 4	5	2	0	−10, 19
1	0	0	−4, 5	6	0	0	−26, 27
2	0	0	−6, 7	7	0	0	−37, 38
3	0	1	2, 5	7	3	0	2, 25
3	1	0	−7, 10	11	1	1	−137, 158
4	0	1	−6, 13	17	2	2	−829, 1270

Proof. Put $\beta = (1 + \sqrt{-7})/2$. Then

$$X^2 - X\cdot Y + 2\cdot Y^2 = (X-\beta\cdot Y)\cdot(X-\bar{\beta}\cdot Y) .$$

Note that $\mathbb{Q}(\sqrt{-7})$ has class number 1 , and that

$$2 = \frac{1 + \sqrt{-7}}{2} \cdot \frac{1 - \sqrt{-7}}{2} , \quad 11 = (2 + \sqrt{-7})\cdot(2 - \sqrt{-7}) .$$

Suppose $\gamma \mid X - \beta\cdot Y$ and $\gamma \mid X - \bar{\beta}\cdot Y$. Then $\gamma \mid (\bar{\beta}-\beta)\cdot Y = - \sqrt{-7}\cdot 3^{m_1}\cdot 7^{m_2}$. On the other hand, $\gamma \mid 11\cdot 2^n$. It follows that $\gamma = \pm 1$, hence $X - \beta\cdot Y$ and $X - \bar{\beta}\cdot Y$ are coprime. Thus we have two possibilities:

$$X - \beta\cdot Y = \pm (2 \pm \sqrt{-7})\cdot\left(\frac{1 \pm \sqrt{-7}}{2} \right)^n ,$$

$$X - \beta\cdot Y = \pm (2 \mp \sqrt{-7})\cdot\left(\frac{1 \pm \sqrt{-7}}{2} \right)^n ,$$

in each equation the 2nd and 3rd $\pm$ being independent. Hence we have to solve

$$G_n^{(j)} = \lambda^{(j)}\cdot\beta^n + \bar{\lambda}^{(j)}\cdot\bar{\beta}^n = 3^{m_1}\cdot 7^{m_2} \quad \text{for} \quad j = 1,\ 2 ,$$

with $G_{n+1}^{(j)} = G_n^{(j)} - 2\cdot G_{n-1}^{(j)}$ for $j = 1,\ 2$, and $\lambda^{(1)} = \bar{\lambda}^{(2)} = (2+\sqrt{-7})/\sqrt{-7}$, so that $G_0^{(1)} = G_0^{(2)} = 1,\ G_1^{(1)} = 3,\ G_1^{(2)} = -1$. Note that $\vartheta_i^{(1)} = -\vartheta_i^{(2)}$ for $i = 1,\ 2$, and $\psi^{(1)} = -\psi^{(2)}$. For $j = 1$ we have solved (4.22) in the example of Section 4.8. It is left to the reader to solve (4.22) for $j = 2$. This can be done with the numerical data given for the case $j = 1$. □

Remark. The computation time for the above proof was less than 3 sec.

Chapter 5. The inequality $0 < x - y < y^{\delta}$ in S-integers.

The results of this chapter have been published in de Weger [1987].

5.1. Introduction.

Let S be the set of all positive integers composed of primes from a fixed finite set $\{ p_1, \ldots, p_s \}$, where $s \geq 2$, and let $\delta \in (0,1)$. In this chapter we study the diophantine inequality

$$0 < x - y < y^{\delta} \tag{5.1}$$

in $x, y \in S$. We give explicit upper bounds for the solutions, and we show how the algorithms for homogeneous, one- and multi-dimensional diophantine approximation in the real case, that were presented in Chapter 3, can be used for finding all solutions of (5.1) for any set of parameters $p_1, \ldots, p_s$, δ. For $s = 2$ the continued fraction method (cf. Section 3.2) is used. For $s \geq 3$ we use the L^3-algorithm for reducing upper bounds (cf. Section 3.7).

Tijdeman [1973] (see also Shorey and Tijdeman [1986], Theorem 1.1) showed that there exists a computable number c, depending on $\max(p_i)$ only, such that for all $x, y \in S$ with $x > y \geq 3$,

$$x - y > y/(\log y)^{c} .$$

Thus, for any solution of (5.1) a bound for x, y follows. Størmer [1897] showed how to solve the equation $x - y = k$ with $k = 1, 2$ with an elementary method (see also Mahler [1935], Lehmer [1964]). Our method can solve this equation for arbitrary $k \in \mathbb{Z}$. For the one-dimensional case $s = 2$, Ellison [1971^b] has proved the following result: for all but finitely many explicitly given exceptions, $| 2^x - 3^y | > \exp\big(x\cdot(\log 2 - 1/10)\big)$ for all $x, y \in \mathbb{N}$. Cijsouw, Korlaar and Tijdeman (appendix to Stroeker and Tijdeman [1982]) have found all the solutions $x, y \in \mathbb{N}$ of the inequality

$$| p^x - q^y | < p^{\delta\cdot x}$$

for all primes p, q with $p < q < 20$, and with $\delta = \frac{1}{2}$. We shall extend these results for many more values of p, q and with $\delta = 0.9$. Further, we determine all the solutions of (5.1) for the multi-dimensional case $s = 6$, $\{p_1, \ldots, p_6\} = \{2, 3, 5, 7, 11, 13\}$ with $\delta = \frac{1}{2}$.

In Section 5.2 we derive upper bounds for the solutions of (5.1). In Sections 5.3 and 5.4 we give a method for reducing such upper bounds in the one- and multi-dimensional cases respectively, and work them out explicitly for some examples. Section 5.5 contains tables with numerical data.

5.2. Upper bounds for the solutions.

We assume that the primes are ordered as $p_1 < \ldots < p_s$. For a solution x, y of (5.1), the finitely many $z \in \mathbb{N}$ for which $z\cdot x, z\cdot y$ is also a solution of (5.1) can be found without any difficulty. Therefore we may assume that $(x,y) = 1$. Put

$$X = \max_{1\leqslant i\leqslant s} \mathrm{ord}_{p_i}(x\cdot y) .$$

Put

$$C_1 = 2^{9\cdot s+26}\cdot s^{s+4}\cdot \max(1,\frac{1}{\log p_1})\cdot \Bigl(\prod_{i=2}^{s} \log p_i\Bigr)\cdot \log(e\cdot \log p_{s-1})/(1-\delta) ,$$

$$C_2 = 2\cdot \log 2/\log p_1 + 2\cdot C_1\cdot \log(e\cdot C_1\cdot \log p_s) .$$

<u>THEOREM 5.1.</u> *The solutions of* (5.1) *satisfy* $X < C_2$.

<u>Proof.</u> If $y \leqslant \frac{1}{2}\cdot x$, then $y^{\delta} > x - y \geqslant y$, which contradicts $y \geqslant 1$. So $y > \frac{1}{2}\cdot x$. Put $\Lambda = \log(x/y)$. Then

$$0 < \Lambda < x/y - 1 < y^{-(1-\delta)} < (\tfrac{1}{2}\cdot x)^{-(1-\delta)} . \tag{5.2}$$

By $x = \max(x,y) \geqslant p_1^X$, we obtain

$$0 < \Lambda < 2^{1-\delta}\cdot p_1^{-(1-\delta)\cdot X} . \tag{5.3}$$

We apply Waldschmidt's result, Lemma 2.4, to Λ, with $n = s$, $q = 2$. Note that the 'independence condition' $[\mathbb{Q}(\sqrt{p_1},\ldots,\sqrt{p_n}):\mathbb{Q}] = 2^n$ holds. Since $p_i \geqslant 3$ we have $V_i = \log p_i$ for $i \geqslant 2$. Thus

$$\Lambda > \exp\left(-(\log X + \log(e\cdot\log p_s))\cdot C_1\cdot(1-\delta)\cdot\log p_1 \right) .$$

Combining this with (5.3) we find

$$X < C_1\cdot\log(e\cdot\log p_s) + \log 2/\log p_1 + C_1\cdot\log X .$$

The result now follows from Lemma 2.1, since $C_1 > e^2$. □

<u>Examples.</u> With $s = 2,\ 2 \leqslant p_i \leqslant 199,\ \delta = 0.9$ we have $C_1 < 2.30\times10^{17}$, $C_2 < 1.97\times10^{19}$.
With $s = 6,\ 2 \leqslant p_i \leqslant 13,\ \delta = \frac{1}{2}$ we find $C_1 < 8.37\times10^{33}$, $C_2 < 1.35\times10^{36}$.

5.3. Reducing the upper bounds in the one-dimensional case.

In this section we work out the examples $s = 2,\ \delta = 0.9$, and p_1, p_2 run through either the set of primes below 200, or the set of non-powers below 50 (we did not use that the p_i are primes). We note that for any other triple p_1, p_2, δ the method works similarly. We prove the following result.

<u>THEOREM 5.2.</u> (a) *The diophantine inequality*

$$| p_1^{x_1} - p_2^{x_2} | < \min\left(p_1^{x_1}, p_2^{x_2} \right)^{\delta} \tag{5.4}$$

with p_1, p_2 *primes such that* $p_1 < p_2 < 200$, *and*

$$\begin{gathered} x_1, x_2 \in \mathbb{Z},\ x_1 \geqslant 2,\ x_2 \geqslant 2, \text{ and either } \delta = \tfrac{1}{2} \\ \text{or } \delta = 0.9,\ \min\left(p_1^{x_1}, p_2^{x_2} \right) > 10^{15} \end{gathered} \tag{5.5}$$

has only the 77 *solutions listed in* Table I.
(b) *The diophantine inequality* (5.4) *with* p_1, p_2 *non-powers such that* $2 \leqslant p_1 < p_2 \leqslant 50$ *and conditions* (5.5), *has only the* 74 *solutions listed in* Table II.

<u>Remarks.</u> The Tables are given in Section 5.5. In Tables I, II the column "delta" gives the real number with

$$| p_1^{x_1} - p_2^{x_2} | = \min\left(p_1^{x_1}, p_2^{x_2} \right)^{\text{delta}} .$$

Note that in Theorem 5.2 we do not demand $(x_1,x_2) = 1$, and in Theorem

5.2(b) we do not demand p_1, p_2 to be primes. The conditions (5.5) are chosen such that the numerous solutions of (5.4) with $\delta = 0.9$ and $\min\left(p_1^{x_1}, p_2^{x_2} \right) \leqslant 10^{15}$ can be found without much effort.

Proof. Write

$$\Lambda = |\ x_1 \cdot \log p_1 - x_2 \cdot \log p_2 \ | , \quad X = \max(x_1, x_2) .$$

We assume that

$$p_1^X > 10^{25} , \tag{5.6}$$

since it is easy to find the remaining solutions. Let $\log p_1/\log p_2$ have the simple continued fraction expansion (cf. Section 3.2)

$$\log p_1/\log p_2 = [\ 0,\ a_1,\ a_2,\ \ldots\ ] ,$$

and let the convergents be r_n/q_n for $n = 1, 2, \ldots$. We may assume that $(x_1, x_2) = 1$. First we show that $x_1 \geqslant x_2$. For if $x_1 < x_2$, then

$$\Lambda = x_2 \cdot \log p_2 - x_1 \cdot \log p_1 > X \cdot \left(\log p_2 - \log p_1 \right) \geqslant X \cdot \log \frac{199}{197} ,$$

and from (5.3) and (5.6) we then infer

$$0.0101 \leqslant 0.0101 \cdot X < X \cdot \log \frac{199}{197} < \Lambda < 2^{0.1} \cdot 10^{-5/2} < 0.0034 ,$$

which is contradictory. Thus $x_1 \geqslant x_2$, hence $X = x_1$. Next we prove that

$$p_1^{X/10} > 3.1 \cdot X . \tag{5.7}$$

Namely, suppose the contrary. Then $2^{X/10} \leqslant 3.1 \cdot X$, and it follows that $X \leqslant 80$. This contradicts $3.1 \cdot X \geqslant p_1^{X/10} > 10^{5/2}$. From (5.3) we infer

$$\left| \frac{x_2}{X} - \frac{\log p_1}{\log p_2} \right| < \frac{2^{0.1}}{\log p_2} \cdot p_1^{-X/10} \cdot \frac{1}{X} . \tag{5.8}$$

It follows from (5.7) that

$$\left| \frac{x_2}{X} - \frac{\log p_1}{\log p_2} \right| < \frac{2^{0.1}}{\log 2} \cdot \frac{1}{3.1 \cdot X^2} < \frac{1}{2 \cdot X^2} .$$

Hence x_2/X is, by Lemma 3.1, a convergent of $\log p_1/\log p_2$, say r_k/q_k . From the example at the end of Section 5.2 we see that $X \leqslant X_0 < 1.97 \times 10^{19}$. We find from (3.7) that $k \leqslant 92.996$, hence $k \leqslant 92$. Lemma 3.1 further yields: if (5.3) holds then

$$a_{k+1} > -2 + p_1^{q_k/10} \cdot \frac{1}{q_k} \cdot \frac{\log p_2}{2^{0.1}} , \tag{5.9}$$

and if

$$a_{k+1} > p_1^{q_k/10} \cdot \frac{1}{q_k} \cdot \frac{\log p_2}{2^{0.1}} \tag{5.10}$$

then (5.3) holds for $(x_1, x_2) = (q_k, r_k)$. We computed the continued fraction expansions and the convergents of all numbers $\log p_1/\log p_2$ in the mentioned ranges for p_1, p_2 exactly up to the index n such that $q_{n-1} \leqslant 1.97\times10^{19} < q_n$ (cf. Section 2.5 for details of the computational method). Note that $n \leqslant 93$. We checked all convergents for (5.9), and subsequently for (5.10). It is possible, though unlikely, that there is a convergent that satisfies (5.9) but fails (5.10). We met only one such a case: $p_1 = 15$, $p_2 = 23$, with $\log 15/\log 23 = [\ 0,\ 1,\ 6,\ 2,\ 1,\ 51,\ \ldots]$, so that $a_5 = 51$, $r_4 = 19$, $q_4 = 22$. Now, (5.9) holds but (5.10) fails, since

$$15^{2.2} \cdot \frac{1}{22} \cdot (\log 19)/2^{0.1} = 51.4\ldots \in [51,53) .$$

We have in this case $\Lambda = 0.002714\ldots < 0.002771\ldots = 2^{0.1}\cdot 15^{-2.2}$, so (5.3) is true. But $\log(15^{22}-23^{19})/\log(23^{19}) = 0.9008\ldots > \delta$, so (5.1) is not true. This example illustrates that (5.3) is weaker than (5.1). Therefore all found solutions of (5.3) have been checked for (5.1) as well. The proof is now completed by the details of the computations, which we omit here. □

Remarks. 1. Theorem 5.2(a) is used in the proof of Theorem 6.2.
2. The computations for the proof of Theorem 5.2 took 35 sec.

5.4. Reducing the upper bounds in the multi-dimensional case.

Now let $s \geqslant 3$. Put $x_i = \mathrm{ord}_{p_i}(x/y)$ for $i = 1, \ldots, s$. Then $X = \max|x_i|$, and

$$\Lambda = \sum_{i=1}^{s} x_i \cdot \log p_i .$$

Note that (5.3) is of the form (3.1). Hence by Theorem 5.1 we can use the method described in Section 3.7 for solving (5.3). We shall do so for the example $s = 6$, $\{ p_1, \ldots, p_6 \} = \{ 2, 3, 5, 7, 11, 13 \}$ (the first six primes), and $\delta = \frac{1}{2}$.

We use small refinements of Lemmas 3.7 and 3.8, devised specially for this application, as follows. Let notation be as in Section 3.7.

LEMMA 5.3. *Let* X_1 *be a positive number such that*

$$\ell(\Gamma) \geqslant \sqrt{\left(4\cdot n^2+(n-1)\cdot\gamma^2\right)}\cdot X_1 \ . \tag{5.11}$$

Then (5.3) *has no solutions with for* $i = 1, \ldots, s$

$$\log\left(\gamma\cdot C\cdot\sqrt{2/s\cdot X_1}\right)/\tfrac{1}{2}\cdot\log p_i \leqslant |x_i| \leqslant X \leqslant X_1 \ . \tag{5.12}$$

LEMMA 5.4. *Suppose that*

$$|\tilde{\Lambda}| > \sum_{i=1}^{s} |x_i| \ . \tag{5.13}$$

Then

$$|x_i| < \log\left(\gamma\cdot C\cdot\sqrt{2/\left(|\lambda|-\sum_{i=1}^{s}|x_i|\right)}\right)/(1-\delta)\cdot\log p_i \ . \tag{5.14}$$

Remark. Lemmas 5.3 and 5.4 are refinements of Lemma 3.8, in that they differentiate between the different x_i . Moreover, Lemma 5.3 has slightly sharper condition and conclusion than Lemma 3.7.

Proofs (of Lemmas 5.3 and 5.4). Analogous to the proofs of Lemmas 3.7 and 3.8, using (5.2) and

$$p_i^{|x_i|} \leqslant \max(x,y) = x < 2\cdot|\Lambda|^{-1/2} \ . \qquad \square$$

THEOREM 5.5. *The diophantine inequality*

$$0 < x - y < \sqrt{y}$$

in $x,\ y \in S = \{\ 2^{x_1}\cdot\ldots\cdot 13^{x_6}\ |\ x_i \in \mathbb{N}_0$ for $i = 1, \ldots, 6\ \}$ *with* $(x,y) = 1$ *has exactly* 605 *solutions. Among those,* 571 *satisfy*

$$\mathrm{ord}_2(x\cdot y) \leqslant 19\ ,\quad \mathrm{ord}_3(x\cdot y) \leqslant 12\ ,\quad \mathrm{ord}_5(x\cdot y) \leqslant 8\ ,$$

$$\mathrm{ord}_7(x\cdot y) \leqslant 7\ ,\quad \mathrm{ord}_{11}(x\cdot y) \leqslant 5\ ,\quad \mathrm{ord}_{13}(x\cdot y) \leqslant 5\ .$$

The remaining 34 *solutions are listed in* Table III.

Remark. The upper bounds for $\mathrm{ord}_{p_i}(x\cdot y)$ given for the 571 solutions not listed in Table III are chosen such that it takes a reasonable amount of computer time to find them all by a brute force method. The list of all 605 solutions is too extensive to be reproduced here.

Proof. By the example at the end of Section 5.2 we know that $X < X_0$ for $X_0 = 1.35\times 10^{36}$. We apply the method described in Section 3.7. Take $C = 10^{240}$ (which is chosen so that it is somewhat larger than X_0^6), and $\gamma = 1$. We applied the L^3-algorithm to the corresponding lattice Γ_1, and found a reduced basis $\underline{c}_1, \ldots, \underline{c}_6$ with $|\underline{c}_1| > 9.40\times 10^{39}$. By Lemma 3.4,

$$\ell(\Gamma_1) > 2^{-5/2}\cdot 9.40\times 10^{39} > 1.66\times 10^{39} .$$

This is larger than $\sqrt{(4\cdot 6^2+5\cdot 1^2)}\cdot X_0 = 1.64\ldots\times 10^{37}$, so (5.11) holds with $X_1 = X_0$. By Lemma 5.3 we find

$$X < \log\left(10^{240}\cdot\sqrt{2/6}\cdot 1.35\times 10^{36}\right)/\tfrac{1}{2}\cdot\log 2 < 1350.4 ,$$

so $X \leqslant 1350$. Next we choose $C = 10^{32}$, $\gamma = 1$, and $X_0 = 1350$. The reduced basis of the corresponding lattice Γ_2 was computed, and we found $|\underline{c}_1| > 2.71\times 10^5$. Hence $\ell(\Gamma_2) > 4.79\times 10^4$, which is larger than $\sqrt{149}\cdot 1350 = 1.64\ldots\times 10^4$. Hence Lemma 5.3 yields for all $i = 1, \ldots, 6$

$$|x_i| < \log\left(10^{32}\cdot\sqrt{2/6}\cdot 1350\right)/\tfrac{1}{2}\cdot\log p_i ,$$

and it follows that

$$\begin{array}{lll} |x_1| \leqslant 187 , & |x_2| \leqslant 118 , & |x_3| \leqslant 80 , \\ |x_4| \leqslant 66 , & |x_5| \leqslant 54 , & |x_6| \leqslant 50 . \end{array} \tag{5.15}$$

Next we choose $C = 10^{12}$, $\gamma = 10^4$. We use Lemma 5.4 as follows. If $|\lambda| > 10^6$ then (5.13) holds by (5.15), and Lemma 5.4 yields

$$\begin{array}{lll} |x_1| \leqslant 67 , & |x_2| \leqslant 42 , & |x_3| \leqslant 29 , \\ |x_4| \leqslant 24 , & |x_5| \leqslant 19 , & |x_6| \leqslant 18 . \end{array} \tag{5.16}$$

All vectors in the corresponding lattice Γ_3 satisfying (5.15) and $|\lambda| < 10^6$ have been computed with the Fincke and Pohst algorithm, cf. Section 3.6. We omit details. We found that there exist only two such vectors, but they do not correspond to solutions of (5.1). Hence all solutions of (5.1) satisfy (5.16). Next, we choose $C = 10^8$, $\gamma = 10^4$. If

$|\lambda| > 5\times10^5$ then Lemma 5.4 yields

$$|x_1| \leqslant 42 , \quad |x_2| \leqslant 27 , \quad |x_3| \leqslant 18 ,$$
$$|x_4| \leqslant 15 , \quad |x_5| \leqslant 12 , \quad |x_6| \leqslant 11 . \tag{5.17}$$

There are 143 vectors in the corresponding lattice Γ_4 satisfying (5.16) and $|\lambda| \leqslant 5\times10^5$. Of them, 2 correspond to solutions of (4.1), namely those with

$$(x_1,\ldots,x_6) = (\ \ 7, \ -5, \ \ 3, \ -9, \ -3, \ \ 8) , \quad \lambda = 257674 ,$$

$$(x_1,\ldots,x_6) = (-10, \ 10, \ -6, \ \ 5, \ -6, \ \ 4) , \quad \lambda = 144817 .$$

Both also satisfy (5.17). Hence all solutions of (5.1) satisfy (5.17). At this point it seems inefficient to choose appropriate parameters C, γ , and a bound for $|\lambda|$ to repeat the procedure with. But the bounds of (5.17) are small enough to admit enumeration. Doing so, we found the result. □

Remark. Theorems 5.2 and 5.5 find applications in solving other exponential diophantine equations, see Stroeker and Tijdeman [1982], Alex [1985[a]], [1985[b]], Tijdeman and Wang [1988], and Section 6.4 of this book.

Remark. The computation of the reduced basis of Γ_1 took 113 sec, where we applied the L^3-algorithm as we described it in Section 3.5, in 12 steps. The direct search for the solutions of (5.17) took 228 sec. The remaining computations (computation of the $\log p_i$ up to 250 decimal digits, of the reduced basis of Γ_2 , and of the short vectors in Γ_3 and Γ_4) took 8 sec. Hence in total we used 349 sec.

5.5. Tables.

Table I. (Theorem 5.2(a)).

p_1	x_1	$p_1^{x_1}$	p_2	x_2	$p_2^{x_2}$	delta
2	3	8	3	2	9	0.00000
3	3	27	5	2	25	0.21534
2	5	32	3	3	27	0.48832
5	3	125	11	2	121	0.28906
2	7	128	11	2	121	0.40575
2	7	128	5	3	125	0.22754
2	8	256	3	5	243	0.46694
7	3	343	19	2	361	0.49512
2	9	512	23	2	529	0.45416
3	7	2187	13	3	2197	0.29941
3	7	2187	47	2	2209	0.40194
13	3	2197	47	2	2209	0.32293
19	3	6859	83	2	6889	0.38504
31	3	29791	173	2	29929	0.47828
2	15	32768	181	2	32761	0.18716
13	7	627 48517	89	4	627 42241	0.48703
2	50	1 12589 99068 42624	47	9	1 11913 04731 02767	0.85259
7	18	1 62841 35979 10449	149	7	1 63043 64614 03549	0.80898
19	12	2 21331 49190 66161	83	8	2 25229 22321 39041	0.88568
2	51	2 25179 98136 85248	19	12	2 21331 49190 66161	0.88532
2	51	2 25179 98136 85248	83	8	2 25229 22321 39041	0.76159
5	22	2 38418 57910 15625	157	7	2 35124 32775 37493	0.87942
13	14	3 93737 63856 99289	89	8	3 93658 88057 02081	0.76282
17	13	9 90457 80329 05937	193	7	9 97473 03260 05057	0.86560
7	19	11 39889 51853 73143	197	7	11 51499 04768 98413	0.87594
61	9	11 69414 60928 34141	197	7	11 51499 04768 98413	0.88743
5	23	11 92092 89550 78125	61	9	11 69414 60928 34141	0.89343
5	23	11 92092 89550 78125	29	11	12 20050 97657 05829	0.89862
29	11	12 20050 97657 05829	199	7	12 35866 42791 61399	0.88268
23	12	21 91462 44320 20321	43	10	21 61148 23132 84249	0.88656
11	16	45 94972 98635 72161	71	9	45 84850 07184 49031	0.84059
5	24	59 60464 47753 90625	73	9	58 87158 67082 67913	0.88642
37	11	177 91762 17794 60413	53	10	174 88747 03655 13049	0.89785
29	12	353 81478 32054 69041	89	9	350 35640 37074 85209	0.88568
23	13	504 03636 19364 67383	163	8	498 31141 43181 21121	0.89040
23	13	504 03636 19364 67383	59	10	511 11675 33006 41401	0.89536
11	17	505 44702 84992 93771	163	8	498 31141 43181 21121	0.89580

11	17	505 44702 84992 93771	23	13	504 03636 19364 67383	0.85578
11	17	505 44702 84992 93771	59	10	511 11675 33006 41401	0.88985
7	21	558 54586 40832 84007	41	11	550 32903 17162 48441	0.89708
19	14	799 00668 57828 84121	31	12	787 66278 37885 49761	0.89710
19	14	799 00668 57828 84121	173	8	802 35917 84760 91681	0.86722
2	60	1152 92150 46068 46976	181	8	1151 93665 78235 00641	0.83013
67	10	1822 83780 45517 61449	107	9	1838 45921 24201 54507	0.88680
47	11	2472 15921 50840 12303	199	8	2459 37419 15531 18401	0.87580
13	17	8650 41591 93813 37933	127	9	8594 75474 86093 97887	0.88441
2	63	9223 37203 68547 75808	53	11	9269 03592 93721 91597	0.87844
3	41	36472 99637 71707 86403	149	9	36197 31987 96201 91349	0.89170
2	65	36893 48814 74191 03232	5	28	37252 90298 46191 40625	0.89721
2	66	73786 97629 48382 06464	97	10	73742 41268 94928 26049	0.83799
3	42	1 09418 98913 15123 59209	101	10	1 10462 21254 11204 51001	0.89916
2	68	2 95147 90517 93528 25856	29	14	2 97558 23267 57994 63481	0.89800
113	10	3 39456 73899 22223 14849	191	9	3 38298 68155 95733 17311	0.87990
53	12	4 91258 90425 67261 54641	199	9	4 89415 46411 90705 61799	0.88284
5	30	9 31322 57461 54785 15625	41	13	9 25103 10231 50136 29321	0.89638
19	17	54 80386 85778 48021 85939	47	13	54 60999 70612 05831 77327	0.88730
23	16	61 32610 41568 09986 48961	151	10	61 62677 95033 67185 14001	0.89400
2	73	94 44732 96573 92904 27392	7	26	93 87480 33764 77543 05649	0.89920
2	75	377 78931 86295 71617 09568	181	10	377 38596 84695 57044 99801	0.86840
2	75	377 78931 86295 71617 09568	41	14	379 29227 19491 55588 02161	0.89368
41	14	379 29227 19491 55588 02161	181	10	377 38596 84695 57044 99801	0.89828
3	49	2392 99329 23061 75295 90083	17	19	2390 72435 68515 13248 47153	0.87071
13	21	2470 64529 07345 03927 04413	89	12	2469 90403 56526 21403 03521	0.84941
103	12	14257 60886 84617 89454 47841	157	11	14285 52404 46318 60195 25093	0.88788
3	51	21536 93963 07555 77663 10747	163	11	21580 60662 62396 00904 07387	0.88933
7	29	32199 05755 81317 97268 37607	13	22	32118 38877 95485 51051 57369	0.89390
11	24	98497 32675 80761 10947 11841	61	14	98768 32533 36131 80951 12441	0.89755
37	16	1 23375 11914 21716 63622 74241	191	11	1 23414 74201 97479 41888 22591	0.86078
2	84	1 93428 13113 83406 67952 98816	199	11	1 93813 41794 57931 33178 02199	0.89319
2	84	1 93428 13113 83406 67952 98816	3	53	1 93832 45667 68001 98967 96723	0.89402
3	53	1 93832 45667 68001 98967 96723	199	11	1 93813 41794 57931 33178 02199	0.84151
7	30	2 25393 40290 69225 80878 63249	31	17	2 25501 16774 16274 31786 82911	0.86903
2	90	123 79400 39285 38027 48991 24224	181	12	123 63541 71303 11583 51179 80561	0.89326
43	17	587 44031 06360 42001 88795 53643	71	15	587 32059 59385 49335 38673 30551	0.86709
2	99	63382 53001 14114 70074 83516 02688	97	15	63325 11891 36789 38604 32759 54593	0.89791
2	102	5 07060 24009 12917 60598 68128 21504	83	16	5 07282 02989 53863 75247 83563 99681	0.89060
13	28	15 50293 28026 62396 21526 95351 05521	89	16	15 49673 14251 78936 43509 93277 30561	0.89106

Table II. (Theorem 5.2(b)).

p_1	x_1	$p_1^{x_1}$	p_2	x_2	$p_2^{x_2}$	delta
2	3	8	3	2	9	.00000
3	3	27	5	2	25	0.21534
2	5	32	3	3	27	0.48832
2	5	32	6	2	36	0.40000
5	3	125	11	2	121	0.28906
2	7	128	11	2	121	0.40575
2	7	128	5	3	125	0.22754
6	3	216	15	2	225	0.40876
2	8	256	3	5	243	0.46694
7	3	343	19	2	361	0.49512
2	9	512	23	2	529	0.45416
2	10	1024	10	3	1000	0.46007
6	4	1296	11	3	1331	0.49607
12	3	1728	42	2	1764	0.48070
2	11	2048	45	2	2025	0.41184
3	7	2187	13	3	2197	0.29941
3	7	2187	47	2	2209	0.40194
13	3	2197	47	2	2209	0.32293
15	4	50625	37	3	50653	0.30762
6	7	2 79936	23	4	2 79841	0.36309
2	50	1 12589 99068 42624	47	9	1 11913 04731 02767	0.85259
2	50	1 12589 99068 42624	18	12	1 15683 13814 26176	0.89628
24	11	1 52168 11431 69024	33	10	1 53157 89852 64449	0.85597
15	13	1 94619 50683 59375	50	9	1 95312 50000 00000	0.83986
2	51	2 25179 98136 85248	19	12	2 21331 49190 66161	0.88532
6	20	3 65615 84400 62976	26	11	3 67034 44869 87776	0.84507
11	15	4 17724 81694 15651	20	12	4 09600 00000 00000	0.89095
28	11	8 29350 94674 71872	39	10	8 14040 60851 91601	0.89154
10	16	10 00000 00000 00000	17	13	9 90457 80329 05937	0.87396
5	23	11 92092 89550 78125	29	11	12 20050 97657 05829	0.89862
2	54	18 01439 85094 81984	30	11	17 71470 00000 00000	0.89096
23	12	21 91462 44320 20321	43	10	21 61148 23132 84249	0.88656
6	21	21 93695 06403 77856	23	12	21 91462 44320 20321	0.81690
6	21	21 93695 06403 77856	43	10	21 61148 23132 84249	0.88845
2	55	36 02879 70189 63968	24	12	36 52034 74360 56576	0.88735

19	13	42 05298 34622 57059	46	10	42 42074 74827 76576	0.87619
3	35	50 03154 50989 99707	33	11	50 54210 65137 26817	0.88076
13	15	51 18589 30140 90757	33	11	50 54210 65137 26817	0.88656
26	12	95 42895 66616 82176	35	11	96 54915 73730 46875	0.88631
35	11	96 54915 73730 46875	50	10	97 65625 00000 00000	0.88575
14	15	155 56809 55578 12224	21	13	154 47237 77391 19461	0.87497
11	17	505 44702 84992 93771	23	13	504 03636 19364 67383	0.85578
7	21	558 54586 40832 84007	41	11	550 32903 17162 48441	0.89708
6	23	789 73022 30536 02816	31	12	787 66278 37885 49761	0.85579
6	23	789 73022 30536 02816	19	14	799 00668 57828 84121	0.89216
19	14	799 00668 57828 84121	31	12	787 66278 37885 49761	0.89710
26	13	2481 15287 32037 36576	47	11	2472 15921 50840 12303	0.86739
28	13	6502 11142 24979 47648	37	12	6582 95200 58400 35281	0.89872
15	16	6568 40835 57128 90625	28	13	6502 11142 24979 47648	0.89414
15	16	6568 40835 57128 90625	37	12	6582 95200 58400 35281	0.85892
2	65	36893 48814 74191 03232	5	28	37252 90298 46191 40625	0.89721
37	13	2 43569 22421 60813 05397	50	12	2 44140 62500 00000 00000	0.87101
2	68	2 95147 90517 93528 25856	29	14	2 97558 23267 57994 63481	0.89800
11	20	6 72749 99493 25600 09201	40	13	6 71088 64000 00000 00000	0.87486
5	30	9 31322 57461 54785 15625	41	13	9 25103 10231 50136 29321	0.89638
35	14	41 39545 12236 93847 65625	46	13	41 29065 87698 35408 01536	0.87993
19	17	54 80386 85778 48021 85939	47	13	54 60990 70612 05831 77327	0.88730
6	28	61 40942 21446 48154 97216	23	16	61 32610 41568 09986 48961	0.86842
2	73	94 44732 96573 92904 27392	7	26	93 87480 33764 77543 05649	0.89920
20	17	131 07200 00000 00000 00000	38	14	130 90925 53986 67734 38464	0.86863
2	74	188 89465 93147 85808 54784	39	14	188 32349 19413 17426 09041	0.88695
2	75	377 78931 86295 71617 09568	41	14	379 29227 19491 55588 02161	0.89368
3	49	2392 99329 23061 75295 90083	17	19	2390 72435 68515 13248 47153	0.87071
15	20	3325 25673 00796 50878 90625	37	15	3334 46267 95181 53070 88493	0.89126
19	19	19784 19655 66031 35891 23979	33	16	19779 85201 46255 88779 34081	0.84943
7	29	32199 05755 81317 97268 37607	13	22	32118 38877 95485 51051 57369	0.89390
2	84	1 93428 13113 83406 67952 98816	3	53	1 93832 45667 68001 98967 96723	0.89402
7	30	2 25393 40290 69225 80878 63249	31	17	2 25501 16774 16274 31786 82911	0.86903
11	25	10 83470 59433 88372 20418 30251	34	17	10 84280 35605 96593 23542 07744	0.87991
14	23	22 95856 92886 98149 54822 20544	18	21	22 94682 51895 12940 71398 72768	0.87516
23	20	171 61558 31334 58634 29238 95201	40	17	171 79869 18400 00000 00000 00000	0.89088
6	35	171 90707 99748 42259 10286 58176	23	20	171 61558 31334 58634 29238 95201	0.89829
6	35	171 90707 99748 42259 10286 58176	40	17	171 79869 18400 00000 00000 00000	0.88250
15	25	25251 16829 40423 48861 69433 59375	43	18	25259 93335 73498 06081 18208 06649	0.88234

Table III . (Theorem 5.5).

x_1	x_2	x_3	x_4	x_5	x_6	x	y	$x-y$
−1	−11	−1	0	6	0	17 71561	17 71470	91
0	4	5	1	−6	0	17 71875	17 71561	314
21	−2	−2	−1	−3	0	20 97152	20 96325	827
1	13	−1	−3	−1	−2	31 88646	31 88185	461
19	0	0	−8	1	0	57 67168	57 64801	2367
6	2	−1	1	−6	3	88 58304	88 57805	499
−2	15	−1	−2	−4	0	143 48907	143 48180	727
11	−15	0	2	1	1	143 50336	143 48907	1429
1	8	−1	−8	0	3	288 29034	288 24005	5029
−22	5	1	−1	1	3	293 62905	293 60128	2777
13	1	3	−1	1	−6	337 92000	337 87663	4337
1	2	9	−4	−4	0	351 56250	351 53041	3209
3	3	0	4	2	−7	627 52536	627 48517	4019
−26	1	0	5	3	0	671 10351	671 08864	1487
3	−13	10	−2	0	0	781 25000	781 21827	3173
8	−2	−10	4	1	1	878 95808	878 90625	5183
25	1	−4	0	−5	0	1006 63296	1006 56875	6421
−6	1	−2	−6	0	7	1882 45551	1882 38400	7151
8	−13	0	3	−2	3	1929 14176	1929 13083	1093
1	−13	−3	7	2	0	1992 97406	1992 90375	7031
−4	−1	−4	1	−4	7	4392 39619	4392 30000	9619
−4	2	−11	2	6	0	7812 58401	7812 50000	8401
16	−3	5	1	−1	−6	14336 00000	14335 62273	37727
−8	8	0	−8	3	2	14758 24779	14757 89056	35723
−5	−2	−5	11	0	−3	19773 26743	19773 00000	26743
−25	7	1	0	−2	5	40600 88955	40600 86272	2683
2	0	13	−9	−2	0	48828 12500	48827 86447	26053
−14	19	−2	−4	1	−1	1 27848 76137	1 27848 44800	31337
−24	−1	−2	12	−1	0	1 38412 87201	1 38412 03200	84001
−5	5	10	0	1	−8	2 61035 15625	2 61033 83072	1 32553
2	−4	−9	3	7	−2	2 67363 98612	2 67363 28125	70487
18	7	0	−13	0	2	9 68892 08832	9 68890 10407	1 98425
7	−5	3	−9	−3	8	1305 16915 36000	1305 16881 72831	33 63169
−10	10	−6	5	−6	4	2834 49801 04623	2834 49760 00000	41 04623

Chapter 6. The equation $x + y = z$ in S-integers.

The results of this chapter have been published in de Weger [1987].

6.1. Introduction.

Let S be the set of all positive integers composed of primes from a fixed finite set $\{ p_1, \ldots, p_s \}$, where $s \geq 3$. This chapter is devoted to the diophantine equation

$$x + y = z \tag{6.1}$$

in $x, y, z \in S$. Without loss of generality we may assume that x, y, z are relatively prime. For any $a \in S$ we define

$$m(a) = \max_{1 \leq i \leq s} \operatorname{ord}_{p_i}(a) .$$

It was proved by Mahler [1933] that (6.1) has only finitely many solutions, but his proof is ineffective. An effective version, i.e. an effectively computable upper bound for $m(x \cdot y \cdot z)$ for the solutions x, y, z of (6.1), can be derived from the results of Coates [1969], [1970] and Sprindžuk [1969], since (6.1) can be reduced to a finite number of Thue equations. See also Chapter 1 of Shorey and Tijdeman [1986].

We derive an explicit upper bound in Section 6.2. Section 6.3 is devoted to some details of the p-adic approximation lattices on which the reduction method of Sections 6.4 and 6.5 are based. In Section 6.4 we give a method of solving (6.1) in the one-dimensional case $s = 3$. This method is based on the reduction procedure given in Section 3.10, and we also use a combination of p-adic and real approximation techniques, similar to that of Section 4.8. But instead of actually performing the real reduction step, we now can simply refer to the results of Chapter 5. As an example we find all the solutions of the slightly more general equation $x \pm y = w \cdot z$, where x, y, z are powers of $2, 3$ or 5, and $w \in \mathbb{Z}$, $|w| \leq 1000000$, $(w,z) = 1$.

In Section 6.5 we give a procedure for solving (6.1) in the multi-dimensional case $s \geqslant 4$, based on the reduction procedure described in Section 3.11. We work out the example $\{ p_1, \ldots, p_6 \} = \{ 2, 3, 5, 7, 11, 13 \}$, and actually determine all the solutions. This generalizes the result of Alex [1976], who gave by elementary arguments a complete solution of (6.1) for the case $\{ p_1, \ldots, p_4 \} = \{ 2, 3, 5, 7 \}$. See also Rumsey and Posner [1964] and Brenner and Foster [1982]. We conclude in Section 6.6 with some remarks on the Oesterlé-Masser conjecture, also known as the 'abc'-conjecture, which is related to equation (6.1). In particular, our method of solving (6.1) leads to a method of finding examples that are of interest with respect to the abc-conjecture. Finally, we give tables in Section 6.7.

6.2. Upper bounds.

We give in this section an upper bound for the solutions of (6.1), based on Lemma 2.6 (cf. Yu [1987]). Note that in de Weger [1987] we used the result of van der Poorten [1977] instead (see also the Correction to de Weger [1987]).

We introduce a lot of notation. Assume that $p_1 < \ldots < p_s$. Let q_i be the smallest prime with $q_i \nmid p_i \cdot (p_i - 1)$ for $i = 1, \ldots, s$. Put

$$t = [2 \cdot s/3] , \quad P = \prod_{i=1}^{s} p_i , \quad q = \max_i q_i ,$$

$C_1(2,t)$ and a_1 as in lemma 2.6 with $n = t$,

$$U = C_1(2,t) \cdot a_1^t \cdot t^{t+5/2} \cdot q^{2 \cdot t} \cdot (q-1) \cdot \log^2(t \cdot q) \cdot \max_i \frac{(p_i - 1) \cdot \left(2 + \frac{1}{p_i - 1}\right)^t}{(\log p_i)^{t+2}} \cdot$$

$$\cdot (\log p_s)^t \cdot \left(\log(4 \cdot \log p_s) + \frac{\log p_s}{8 \cdot t} \right) ,$$

$$C_1 = U/6 \cdot t , \quad C_2 = U \cdot \log 4 ,$$

$$V_i = \max(1, \log p_i) \quad \text{for} \quad i = s-t+1, \ldots, s , \quad \Omega = \prod_{i=s-t+1}^{s} V_i ,$$

$$C_3 = 2^{9 \cdot t+26} \cdot t^{t+4} \cdot \Omega \cdot \log(e \cdot V_{s-1}) ,$$

$$C_4 = \max \left(7.4, \left(C_1 \cdot \log(P/p_1) + C_3\right) / \log p_1 \right) ,$$

$$C_5 = \left(C_2 \cdot \log(P/p_1) + C_3 \cdot \log(e \cdot V_s) + 0.327\right) / \log p_1 ,$$

$$C_6 = \max \left(C_5,\ (C_2 \cdot \log(P/p_1) + \log 2)/\log p_1 \right) ,$$

$$C_7 = 2 \cdot \left(C_6 + C_4 \cdot \log C_4 \right) ,$$

$$C_8 = \max \left(p_s,\ \log\left(2 \cdot (P/p_1)^{p_s}\right)/\log p_1,\ C_2 + C_1 \cdot \log C_7,\ C_7 \right) .$$

Now we state the main result.

<u>THEOREM 6.1.</u> *The solutions of* (6.1) *satisfy* $m(x \cdot y \cdot z) \leqslant C_8$.

<u>Proof.</u> If we consider instead of (6.1) the equivalent equation

$$x \pm y = z \tag{6.2}$$

then we may assume that $x \cdot y$ has at most t prime divisors, $p_{i_1}, \ldots, p_{i_t}$ say. Suppose first that $m(x \cdot y) \leqslant p_s$. Then

$$p_1^{m(z)} \leqslant z \leqslant 2 \cdot \max(x,y) < 2 \cdot (P/p_1)^{p_s} ,$$

hence

$$m(x \cdot y \cdot z) < \max \left(p_s,\ \log\left(2 \cdot (P/p_1)^{p_s}\right)/\log p_1 \right) \leqslant C_8 .$$

Next suppose that $m(x \cdot y) \geqslant p_s$ and $m(z) \geqslant 2$. Then for some $p = p_i$,

$$m(z) = \mathrm{ord}_p(z) = \mathrm{ord}_p \left\| \pm\frac{x}{y} - 1 \right\| = \mathrm{ord}_p \left\| \log_p\left(\frac{x}{y}\right) \right\| .$$

Put $x/y = \prod_{j=1}^{t} p_{i_j}^{x_{i_j}}$. Then $m(x \cdot y) = \max_{1 \leqslant j \leqslant t} |x_{i_j}|$. We apply Lemma 2.6 (Yu's lemma) with $n = t$, $B_0 = B_n = B' = B = m(x \cdot y)$. Since $m(x \cdot y) \geqslant p_s$ and $t \geqslant 2$ we have

$$W = \max \left\| \log\left(1+\frac{3}{4 \cdot t} \cdot B\right),\ \log B,\ \log p \right\| = \log B .$$

Note that $C_1(p,n)$ is maximal for $p = 2$. We obtain

$$m(z) < C_1 \cdot \log m(x \cdot y) + C_2 . \tag{6.3}$$

Obviously (6.3) is true if $m(z) < 2$. If in (6.2) the plus sign holds, then

$$(P/p_1)^{m(z)} \geqslant z > \max(x,y) \geqslant p_1^{m(x \cdot y)} .$$

By (6.3) and $C_3 > 0$ it then follows that

$$m(x\cdot y) < C_4\cdot \log m(x\cdot y) + C_6 \ . \tag{6.4}$$

Next suppose that in (6.2) the minus sign holds. Then we apply Lemma 2.4 to prove (6.4) for this case, as follows. Suppose (6.4) is false. Then

$$\left| \frac{y}{x} - 1 \right| = \frac{z}{x} = \frac{z}{\max(x,y)} \leqslant \frac{(P/p_1)^{m(z)}}{p_1^{m(x\cdot y)}} < \frac{(P/p_1)^{C_1\cdot\log m(x\cdot y)+C_2}}{p_1^{C_4\cdot\log m(x\cdot y)+C_6}} ,$$

which is less than $\frac{1}{2}$, by the definition of C_4 and C_6 . Hence

$$\left|\log \frac{y}{x}\right| < (2\cdot\log 2)\cdot\left| \frac{y}{x} - 1 \right| < (2\cdot\log 2)\cdot\frac{(P/p_1)^{C_1\cdot\log m(x\cdot y)+C_2}}{p_1^{m(x\cdot y)}} .$$

On the other hand, Lemma 2.4 yields

$$\left|\log \frac{y}{x}\right| > \exp\left(-C_3\cdot(\log m(x\cdot y) + \log(e\cdot V_s)) \right) .$$

Thus we obtain

$$m(x\cdot y)\cdot\log p_1 < \log(2\cdot\log 2) + (C_1\cdot\log m(x\cdot y)+C_2)\cdot\log(P/p_1)$$

$$+ C_3\cdot(\log m(x\cdot y) + \log(e\cdot V_s)) \leqslant (\log p_1)\cdot(C_4\cdot\log m(x\cdot y)+C_6) .$$

This contradicts our assumption that (6.4) if false. Consequently (6.4) is true in all cases. Now, by $C_4 > e^2$, Lemma 2.1 yields $m(x\cdot y) < C_7$, and (6.3) then yields $m(x\cdot y\cdot z) < C_8$. □

<u>Examples.</u> If $s = 3$, $\{ p_1, p_2, p_3 \} = \{ 2, 3, 5 \}$ then $C_8 < 3.98\times10^{17}$. If $s = 6$, $\{ p_1, \ldots, p_6 \} = \{ 2, 3, 5, 7, 11, 13 \}$ then $C_8 < 5.60\times10^{27}$.

6.3. The p-adic approximation lattices.

As in the proof of Theorem 6.1 we consider (6.2) instead of (6.1). Let p be any of the primes $p_1, \ldots, p_s$. We may assume that $p \nmid x\cdot y$. Rename the other primes as $p_0, \ldots, p_{s-2}$, such that $\mathrm{ord}_p(\log_p(p_0))$ is minimal. For $i = 1, \ldots, s-2$ put (cf. Section 3.11)

$$\vartheta_i = -\log_p(p_i)/\log_p(p_0) = \sum_{\ell=0}^{\infty} u_{i,\ell}\cdot p^{\ell} ,$$

where $u_{i,\ell} \in \{ 0, 1, \dots, p-1 \}$. The ϑ_i take the place of the ϑ'_i of Section 3.11. Then it is clear from Section 3.11 how to define the p-adic approximation lattices Γ_μ for $\mu \in \mathbb{N}_0$. Put

$$\Lambda = \sum_{i=1}^{s-2} x_i \cdot \vartheta_i - x_0 .$$

Then Lemma 3.13 yields

$$\Gamma_\mu = \langle (x_1, \dots, x_{s-2}, x_0) \mid |\Lambda|_p \leqslant p^{-\mu} \rangle$$

$$= \langle (x_1, \dots, x_{s-2}, x_0) \mid \left| \log_p \left(\prod_{i=0}^{s-2} p_i^{x_i} \right) \right|_p \leqslant p^{-(\mu+\mu_0)} \rangle ,$$

where $\mu_0 = \mathrm{ord}_p(\log_p(p_0))$. In Section 3.13 we studied the set

$$\Gamma^*_\mu = \langle (x_1, \dots, x_{s-2}, x_0) \mid \left| \prod_{i=0}^{s-2} p_i^{x_i} \pm 1 \right|_p \leqslant p^{-(\mu+\mu_0)} \rangle ,$$

which is a sublattice of Γ_μ . In Lemma 3.17 we showed how a basis of Γ^*_μ can be found from a basis of Γ_μ . In practice this is very easy, especially if for $p \geqslant 5$ it happens to be possible to choose p_0 such that not only $\mathrm{ord}_p(\log_p(p_0))$ is minimal, but also p_0 is a primitive root (mod p) . Then, using the notation of Lemma 3.17 (with $\underline{b}_0$ as the last element of the basis), choose $\zeta \equiv p_0 \pmod p$. Then $k(\underline{b}_0) = 1$, and it follows that $\underline{b}'_i = \underline{b}_i$ for $i = 1, \dots, s-2$. By $\underline{b}_i = \left(0, \dots, 1, \dots, 0, \vartheta_i^{(\mu)}\right)^T$ we have

$$p_i \cdot p_0^{\vartheta_i^{(\mu)}} \equiv \zeta^{k(\underline{b}_i)} \pmod{p^{\mu+\mu_0}} .$$

If $p_i \equiv p_0^{\alpha_i} \pmod p$, then it follows that

$$\gamma^*_i \equiv \alpha_i + \vartheta_i^{(\mu)} \equiv \alpha_i + \sum_{\ell=0}^{\mu-1} u_{i,\ell} \pmod{(p-1)/2} \quad \text{for } i = 1, \dots, s-2 ,$$

$$\gamma^*_0 = (p-1)/2 .$$

Lemma 3.14 (with $c_1 = 0$, $c_2 = 1$) now yields: if

$$\ell(\Gamma^*_\mu) > \sqrt{(s-1)} \cdot X_1 \tag{6.5}$$

then (6.2) has no solutions with

$$\mu + \mu_0 \leqslant \mathrm{ord}_p(z) \leqslant m(x \cdot y \cdot z) \leqslant X_1 . \tag{6.6}$$

6.4. Reducing the upper bounds in the one-dimensional case.

In Section 3.10 we have described how an upper bound for the solutions of (6.1) in the case $s = 3$ can be reduced. We shall apply that method in this section to the following problem.

THEOREM 6.2. *The diophantine equation*

$$x \pm y = w \cdot z \,, \tag{6.7}$$

where $x = p_0^{x_0}$, $y = p_1^{x_1}$, $z = p^u$, $(p,p_0,p_1) = (2,3,5),\ (3,2,5),\ (5,2,3)$, $x_0,\ x_1,\ u \in \mathbb{N}_0$, $w \in \mathbb{Z}$, $|w| \leqslant 10^6$, *and* $p \nmid w$, *has exactly* 291 *solutions for* $p = 2$, 412 *solutions for* $p = 3$, *and* 570 *solutions for* $p = 5$. *In* Table I *all solutions with* $u \geqslant 3$ *are given. The solutions with* $u \leqslant 2$ *satisfy* $x_0 \leqslant 14$, $x_1 \leqslant 9$ *for* $p = 2$, $x_0 \leqslant 23$, $x_1 \leqslant 10$ *for* $p = 3$, *and* $x_0 \leqslant 25$, $x_1 \leqslant 15$ *for* $p = 5$.

Remark. It is easy to find all solutions of (6.7) with $u \leqslant 2$. The Tables are presented in Section 6.7.

Proof. Put $X = \max\limits_{p=2,3,5} \mathrm{ord}_p(x \cdot y \cdot z)$. The example at the end of Section 6.2 shows that in the case $|w| = 1$ we have $X < 3.98 \times 10^{17}$. It can be checked without difficulties that the effect of the w with $|w| \leqslant 10^6$ in the proof of Theorem 6.1 can be neclected (it disappears in the rounding off), so that for the solutions of (6.7) also $X < X_0 = 3.98 \times 10^{17}$ holds. Put

$$x/y = p_0^{y_0} \cdot p_1^{y_1} \,, \quad \vartheta = -\log_p(p_1)/\log_p(p_0) \,.$$

Note that ϑ is a p-adic integer. Define the lattices Γ_μ, Γ_μ^* as in Section 6.3, so Γ_μ is generated by

$$\underline{b}_1 = \begin{pmatrix} 1 \\ \vartheta^{(\mu)} \end{pmatrix}, \quad \underline{b}_0 = \begin{pmatrix} 0 \\ p^\mu \end{pmatrix}.$$

For $p = 2,\ 3$ we have $\Gamma_\mu^* = \Gamma_\mu$, and for $p = 5$ a basis of Γ_μ^* is

$$\underline{b}_1^* = \underline{b}_1 - \gamma \cdot \underline{b}_0 \,, \quad \underline{b}_0^* = 2 \cdot \underline{b}_0 \,,$$

where $\gamma = 0$ if $\vartheta^{(\mu)}$ is odd, $\gamma = 1$ if $\vartheta^{(\mu)}$ is even. Using the algorithm given in Section 3.10, Fig. 3, we can compute a basis $\underline{c}_1$, $\underline{c}_2$ of Γ_μ^* that is reduced in the sense that $|\underline{c}_1| = \ell(\Gamma_\mu^*)$. We did so, with μ as in the following table.

p	p_0	p_1	μ_0	μ	γ	$\|\underline{c}_1\| >$	$u \leqslant$	W	$\|y_0\| \leqslant$	$\|y_1\| \leqslant$
2	3	5	2	143	-	2.68×10^{21}	144	$10^6\cdot2^{144}$	114	78
3	2	5	1	91	-	2.32×10^{21}	91	$10^6\cdot3^{91}$	182	78
5	2	3	1	65	0	5.28×10^{22}	65	$10^6\cdot5^{65}$	189	119

The values of $\vartheta^{(\mu)}$ can be found in Table III. Making an exception to our policy, we give the reduced bases of the Γ_μ^* below (in base p notation):

$p = 2$:

$$\underline{c}_1 = \begin{pmatrix} \text{10 00000 00100 10001 10110 01110 01101} \\ \text{00001 11101 00101 00100 11100 01111 11010 00011} \\ \text{- 1 00010 00110 01000 01011 01110 00010} \\ \text{00101 11000 00000 11100 01111 01011 10111 00001} \end{pmatrix},$$

$$\underline{c}_2 = \begin{pmatrix} \text{10 11011 10000 01011 01101 11000 00111} \\ \text{11001 10100 11011 00000 11111 10110 10110 00001} \\ \text{10 01110 11101 10111 11000 00100 10101} \\ \text{00111 00001 10101 00110 10011 00111 00101 10101} \end{pmatrix},$$

$p = 3$:

$$\underline{c}_1 = \begin{pmatrix} \text{- 102 01121 02221 00210 12120 20020 22222 10212 20222} \\ \text{21002 00122 21100 11102 22102 20001 11222 02212 21011} \end{pmatrix},$$

$$\underline{c}_2 = \begin{pmatrix} \text{-10 12210 12111 01102 02010 12112 12210 21122 21011 20102} \\ \text{- 2 22021 11012 01000 12021 00211 12221 22121 21220 12122} \end{pmatrix},$$

$p = 5$:

$$\underline{c}_1 = \begin{pmatrix} \text{- 211 32230 21042 22023 30141 33034 21420} \\ \text{- 22104 43102 43111 03114 30134 23410} \end{pmatrix},$$

$$\underline{c}_2 = \begin{pmatrix} \text{340 34003 02404 12120 03412 22030 32211} \\ \text{- 414 20001 42202 42210 34043 20120 00432} \end{pmatrix}.$$

From this we found the lower bounds for $|\underline{c}_1|$ given above. They are all larger than $\sqrt{2}\cdot3.98\times10^{17}$. Hence (6.5) holds for $X_1 = X_0$, and then we infer from (6.6) that $u \leqslant \mu + \mu_0 - 1$, and $|w|\cdot z \leqslant W$ as shown in the table above. We now find the new upper bounds for $|y_0|$, $|y_1|$ as follows. If in (6.7) the minus sign holds, supposing that $\min(x,y) > W^{10/9}$, we infer

$$|\ x - y\ | = |w|\cdot z \leqslant W < \min(x,y)^{0.9}.$$

By Theorem 5.2(a), the inequality $|x - y| < \min(x,y)^{0.9}$ has no solutions with $\min(x,y) > W$, since $W > 10^{49}$. Hence $\min(x,y) \leqslant W^{10/9}$, and thus

$$\max(x,y) \leqslant \min(x,y) + |w| \cdot z \leqslant W^{10/9} + W .$$

If in (6.7) the plussign holds, then this inequality follows at once. So now the bounds given in the above table for $|y_0|$, $|y_1|$ follow from

$$|y_i| \cdot \log p_i \leqslant \log \max(x,y) \leqslant \log(W^{10/9}+W) .$$

We repeat the procedure with μ as in the following table.

p	μ	γ	$\lvert\underline{c}_1\rvert >$	$\sqrt{2}\cdot X_0 <$	$u \leqslant$	W	$\lvert y_0\rvert \leqslant$	$\lvert y_1\rvert \leqslant$
2	16	-	167.7	161.3	17	$10^6 \cdot 2^{17}$	31	21
3	13	-	535.8	257.4	13	$10^6 \cdot 3^{13}$	49	21
5	7	1	276.1	267.3	7	$10^6 \cdot 5^7$	49	31

The numbers are now so small that the computations can be performed by hand. For example, for $p = 5$, the lattice Γ_7^* is generated by

$$\underline{b}_1^* = \begin{pmatrix} 1 \\ -45607 \end{pmatrix}, \quad \underline{b}_0^* = \begin{pmatrix} 0 \\ 156250 \end{pmatrix},$$

and a reduced basis is

$$\underline{c}_1 = \begin{pmatrix} 185 \\ 205 \end{pmatrix}, \quad \underline{c}_0 = \begin{pmatrix} -394 \\ 408 \end{pmatrix}.$$

We find upper bounds for u and W as given in the above table. In all three cases, $W^{10/9} < 10^{15}$. On supposing $\min(x,y) > 10^{15}$ we infer

$$|x - y| = |w| \cdot z \leqslant W < 10^{15 \cdot 0.9} \leqslant \min(x,y)^{0.9} .$$

By Theorem 5.2(a) we see that the inequality $|x - y| < \min(x,y)^{0.9}$ has only two solutions: $(x,y) = (2^{65}, 5^{28})$, $(2^{84}, 3^{53})$. However, both have $|x - y| > 10^{15 \cdot 0.9}$. So we infer $\min(x,y) \leqslant 10^{15}$, hence by $\max(x,y) \leqslant 10^{15} + W$ we obtain the bounds for $|y_0|$, $|y_1|$ as given above. These bounds are small enough to admit enumeration of the remaining cases. □

Remark. The computer calculations for the above proof took less than 1 sec.

6.5. Reducing the upper bounds in the multi-dimensional case.

In Section 3.11 we have described how an upper bound for the solutions of (6.1) in the case $s \geqslant 3$ can be reduced. We shall apply that method in this section to the following problem.

THEOREM 6.3. *The diophantine equation*

$$x + y = z \tag{6.8}$$

in $x,\ y,\ z \in S = \langle\, 2^{x_1} \cdot \ldots \cdot 13^{x_6} \mid x_i \in \mathbb{N}_0 \ \text{ for } \ i = 1, \ldots, 6 \,\rangle$ *with* $(x,y) = 1$ *and* $x \leqslant y$ *has exactly* 545 *solutions. Of them,* 514 *satisfy*

$$\mathrm{ord}_2(x\cdot y\cdot z) \leqslant 12\ ,\quad \mathrm{ord}_3(x\cdot y\cdot z) \leqslant 7\ ,\quad \mathrm{ord}_5(x\cdot y\cdot z) \leqslant 5\ ,$$

$$\mathrm{ord}_7(x\cdot y\cdot z) \leqslant 4\ ,\quad \mathrm{ord}_{11}(x\cdot y\cdot z) \leqslant 3\ ,\quad \mathrm{ord}_{13}(x\cdot y\cdot z) \leqslant 3\ .$$

The remaining 31 *solutions are given in* Table II.

Remark. From Theorem 6.3 it is easy to compute all 545 solutions of (6.8).

Proof. In the example at the end of Section 6.2 we have seen that $m(x\cdot y\cdot z) < X_0 = 5.60\times10^{27}$. With the notation of Section 6.3 we choose the following parameters.

p	p_0	p_1	p_2	p_3	p_4	μ_0	μ	γ_0^*	γ_1^*	γ_2^*	γ_3^*	γ_4^*
2	3	5	7	11	13	2	605	-	-	-	-	-
3	2	5	7	11	13	1	385	-	-	-	-	-
5	2	3	7	11	13	1	275	2	0	1	1	1
7	3	2	5	11	13	1	220	3	0	-1	-1	0
11	2	3	5	7	13	1	165	5	2	0	-1	-1
13	2	3	5	7	11	1	165	6	-2	-1	-2	3

We computed the six values of the $\vartheta_i^{(\mu)}$ for $i = 1,\ 2,\ 3,\ 4$ (and give them in Table III), and the reduced bases of the six lattices Γ_μ^* , by the L^3-algorithm. Thus we obtained lower bounds for $\ell(\Gamma_\mu^*)$ as in the following table. They are all larger than $\sqrt{5}\cdot 5.60\times10^{27}$ (note that we have a very large margin here, we could have taken the μ's probably about 20% smaller). So we apply Lemma 3.14 for $X_1 = X_0 = 5.60\times10^{27}$. For every p we thus find $\mathrm{ord}_p(z) \leqslant \mu + \mu_0 - 1$. Since (6.2) is invariant under permutations of $x,\ y,\ z$, we even have $\mathrm{ord}_p(x\cdot y\cdot z) \leqslant \mu + \mu_0 - 1$, as shown in the next table.

p	$\ell(\Gamma_\mu^*) \geqslant \lvert \underline{c}_1 \rvert /4 >$	$\mathrm{ord}_p(x\cdot y\cdot z) \leqslant$
2	4.70×10^{35}	606
3	1.15×10^{36}	385
5	6.27×10^{37}	275
7	3.17×10^{36}	220
11	5.74×10^{33}	165
13	1.73×10^{36}	165

Hence $m(x\cdot y\cdot z) \leqslant 606$.

We repeated the procedure with $X_0 = 606$ and μ as in the following table. After computing the reduced bases of the six lattices Γ_μ^* we found the following data. Note that in all cases $\ell(\Gamma_\mu^*) \geqslant \sqrt{5}\cdot 606$.

p	μ	γ_0^*	γ_1^*	γ_2^*	γ_3^*	γ_4^*	$\ell(\Gamma_\mu^*) >$	$\mathrm{ord}_p(x\cdot y\cdot z) \leqslant$
2	66	-	-	-	-	-	1909	67
3	42	-	-	-	-	-	2304	42
5	30	2	0	0	1	1	3417	30
7	24	3	-1	0	1	-1	2391	24
11	18	5	0	-2	2	-1	1443	18
13	18	6	0	1	1	-2	3196	18

Hence $m(x\cdot y\cdot z) \leqslant 67$. Next, we repeated the procedure with $X_0 = 67$, and μ as in the following table. We found

p	μ	γ_0^*	γ_1^*	γ_2^*	γ_3^*	γ_4^*	$\ell(\Gamma_\mu^*) >$	$\mathrm{ord}_p(x\cdot y\cdot z) \leqslant$
2	55	-	-	-	-	-	364	56
3	35	-	-	-	-	-	301	35
5	25	2	1	1	1	0	622	25
7	20	3	-1	1	-1	0	693	20
11	15	5	-1	-2	2	2	192	15
13	15	6	-1	0	3	-2	658	15

Hence $m(x\cdot y\cdot z) \leqslant 56$.

To find the solutions of (6.2) with $\mathrm{ord}_p(x\cdot y\cdot z)$ below the bounds given in the above table, we followed the following procedure. Suppose that we are at a certain moment interested in finding the solutions with $\mathrm{ord}_p(x\cdot y\cdot z) \leqslant f(p)$ where $f(p)$ is given for $p = 2, \ldots, 13$. Choose p , and $\mu < f(p) - \mu_0$, and consider the lattice Γ_μ^* for these values of p, μ. If a solution x, y, z of (6.2) exists with $\mathrm{ord}_p(z) \geqslant \mu + \mu_0$, then the vector $(x_1,\ldots,x_4,x_0)^T$ with $x_i = \mathrm{ord}_{p_i}(x/y)$ for $i = 0, \ldots, 4$, is in the lattice. Its length is bounded by $\sqrt{(f(p_0)^2+\ldots+f(p_4)^2)}$. All vectors in Γ_μ^* with length below this bound can be computed by the algorithm of Fincke and Pohst, as given in Section 3.6. Then all solutions of (6.2) corresponding to lattice points can be selected. Then we replace $f(p)$ by $\mu + \mu_0 - 1$, and we repeat the procedure for newly chosen p, μ.

We performed this procedure, starting with the bounds for $\mathrm{ord}_p(x \cdot y \cdot z)$ given in the above table for $f(p)$, and with p, m as in Table IV (where # stands for the number of solutions of (5.2) found at that stage). At the end we have $f(2) = 4$, $f(p) = 1$ for $p = 3, \ldots, 13$. The remaining solutions can be found by hand. □

Remarks. 1. Theorems 6.2 and 6.3 have applications in group theory (cf. Alex [1976]). We use Theorem 6.3 in Section 7.2.

2. The computer calculations for the proof of Theorem 6.3 took 438 sec., of which 412 were used for the first reduction step. In this first step we applied the L^3-algorithm in 11 steps (cf. Section 3.5), which cost on average about 60 sec. per lattice. The remaining 50 sec. were mainly used for the computation of the 24 $\vartheta_i^{(\mu)}$'s.

6.6. Examples related to the abc-conjecture.

Let x, y, z be positive integers. Put

$$G = \prod_{\substack{p \mid xyz \\ p \text{ prime}}} p .$$

For all x, y, z with $(x,y) = 1$ and $x + y = z$ we define

$$c(x,y,z) = \log z / \log G$$

(called the *Masser-ratio*, according to Tijdeman [1989]). Recently, Oesterlé posed the problem to decide whether there exists an absolute constant C such that $c(x,y,z) < C$ for all x, y, z. Masser [1985] conjectured the stronger assertion that $c(x,y,z) < 1 + \varepsilon$, when z exceeds some bound depending on ε only, for all $\varepsilon > 0$. For a survey of related results and conjectures, see Stewart and Tijdeman [1986], Vojta [1987], Tijdeman [1989].

It might be interesting to have some empirical results on $c(x,y,z)$, and to search for x, y, z for which it is large. From the preceding sections it may be clear that such x, y, z correspond to relatively short vectors in appropriate p-adic approximation lattices.

As a byproduct of the proofs of Theorems 5.5 and 6.3 we computed the value of $c(x,y,z)$, corresponding to many short vectors that we came across in performing the algorithm of Fincke and Pohst. All examples that we found with $c(x,y,z) \geqslant 1.4$ are listed below. Our search was rather unsystematic, so we do not guarantee that this list is complete in any sense.

x	y	z	c(x,y,z)
11^2	$3^2 \cdot 5^6 \cdot 7^3$	$2^{21} \cdot 23$	1.62599
1	$2 \cdot 3^7$	$5^4 \cdot 7$	1.56789
7^3	3^{10}	$2^{11} \cdot 29$	1.54708
$5^2 \cdot 7937$	7^{13}	$2^{18} \cdot 3^7 \cdot 13^2$	1.49762
11^2	$3^9 \cdot 13$	$2^{11} \cdot 5^3$	1.48887
37	2^{15}	$3^8 \cdot 5$	1.48291
$2^7 \cdot 5^2$	$7^6 \cdot 41$	13^6	1.46192
1	$2^5 \cdot 3 \cdot 5^2$	7^4	1.45567
$2^{19} \cdot 13 \cdot 103$	7^{11}	$3^{11} \cdot 5^3 \cdot 11^2$	1.45261
1	$2^{12} \cdot 5^3$	$3^5 \cdot 7^2 \cdot 43$	1.44331
1	$2^4 \cdot 3^7 \cdot 547$	$5^8 \cdot 7^2$	1.43906
$2^{10} \cdot 7$	5^7	$3^8 \cdot 13$	1.43501
3	5^3	2^7	1.42657
5	3^{11}	$2^{10} \cdot 173$	1.41268

Two more examples with $c(x,y,z) \geqslant 1.4$ are known:

$$x = 1\ , \quad y = 3 \cdot 5 \cdot 47^2\ , \quad z = 2^{18} \cdot 79\ , \quad c(x,y,z) = 1.44965\ ,$$

found by G. Frey (communicated to us by Prof. F. Oort), and

$$x = 2\ , \quad y = 109 \cdot 3^{10}\ , \quad z = 23^5\ , \quad c(x,y,z) = 1.62991\ ,$$

found by E. Reyssat (communicated to us by Prof. M. Waldschmidt), which wins the race. Note that these two examples show large primes at two places.

These results do not seem to yield any heuristical evidence for the truth or falsity of the abc-conjecture.

6.7. Tables.

Table I. (Theorem 6.2.)

$p=2,\ p_0=3,\ p_1=5$

x_0	$p_0^{x_0}$	x_1	$p_1^{x_1}$	sign	u	w
2	9	10	9765625	−1	4	−610351
10	59049	10	9765625	−1	4	−606661
4	81	12	244140625	−1	9	−476837
6	729	10	9765625	−1	5	−305153
2	9	8	390625	−1	3	−48827
6	729	8	390625	−1	3	−48737
10	59049	8	390625	−1	3	−41447
14	4782969	10	9765625	−1	7	−38927
4	81	8	390625	−1	4	−24409
0	1	8	390625	−1	5	−12207
8	6561	8	390625	−1	6	−6001
0	1	6	15625	−1	3	−1953
4	81	6	15625	−1	3	−1943
8	6561	6	15625	−1	3	−1133
6	729	6	15625	−1	4	−931
2	9	4	625	−1	3	−77
2	9	6	15625	−1	8	−61
0	1	4	625	−1	4	−39
4	81	4	625	−1	5	−17
0	1	2	25	−1	3	−3
2	9	2	25	−1	4	−1
1	3	1	5	1	3	1
1	3	3	125	1	7	1
2	9	0	1	−1	3	1
3	27	1	5	1	5	1
4	81	0	1	−1	4	5
4	81	2	25	−1	3	7
6	729	2	25	−1	6	11
6	729	4	625	−1	3	13
3	27	3	125	1	3	19
5	243	3	125	1	4	23
5	243	1	5	1	3	31
7	2187	5	3125	1	6	83
6	729	0	1	−1	3	91
7	2187	1	5	1	4	137
11	177147	1	5	1	10	173
3	27	5	3125	1	4	197
8	6561	0	1	−1	5	205
7	2187	3	125	1	3	289
8	6561	4	625	−1	4	371

Table continued

Table I. (cont.)

x_0	$p_0^{x_0}$	x_1	$p_1^{x_1}$	sign	u	w
1	3	5	3125	1	3	391
5	243	5	3125	1	3	421
9	19683	3	125	1	5	619
8	6561	2	25	−1	3	817
10	59049	6	15625	−1	5	1357
5	243	7	78125	1	5	2449
9	19683	1	5	1	3	2461
9	19683	5	3125	1	3	2851
10	59049	2	25	−1	4	3689
12	531441	4	625	−1	7	4147
1	3	7	78125	1	4	4883
9	19683	7	78125	1	4	6113
13	1594323	7	78125	1	8	6533
10	59049	4	625	−1	3	7303
10	59049	0	1	−1	3	7381
12	531441	8	390625	−1	4	8801
3	27	7	78125	1	3	9769
7	2187	7	78125	1	3	10039
11	177147	5	3125	1	4	11267
3	27	9	1953125	1	7	15259
11	177147	3	125	1	3	22159
11	177147	7	78125	1	3	31909
12	531441	0	1	−1	4	33215
12	531441	6	15625	−1	3	64477
12	531441	2	25	−1	3	66427
11	177147	9	1953125	1	5	66571
13	1594323	3	125	1	4	99653
7	2187	9	1953125	1	4	122207
14	4782969	2	25	−1	5	149467
13	1594323	1	5	1	3	199291
13	1594323	5	3125	1	3	199681
1	3	9	1953125	1	3	244141
5	243	9	1953125	1	3	244171
9	19683	9	1953125	1	3	246601
14	4782969	6	15625	−1	4	297959
13	1594323	9	1953125	1	3	443431
15	14348907	5	3125	1	5	448501
14	4782969	8	390625	−1	3	549043
14	4782969	4	625	−1	3	597793
14	4782969	0	1	−1	3	597871
16	43046721	0	1	−1	6	672605
9	19683	11	48828125	1	6	763247
15	14348907	1	5	1	4	896807

Table continued

Table I. (cont.)

$p=3,\ p_0=2,\ p_1=5$

x_0	$p_0^{x_0}$	x_1	$p_1^{x_1}$	sign	u	w
14	16384	10	9765625	−1	4	−120361
9	512	9	1953125	−1	3	−72319
4	16	8	390625	−1	3	−14467
12	4096	6	15625	−1	3	−427
7	128	5	3125	−1	4	−37
2	4	4	625	−1	3	−23
1	2	2	25	1	3	1
5	32	1	5	−1	3	1
6	64	3	125	1	3	7
11	2048	4	625	1	5	11
9	512	0	1	1	3	19
10	1024	2	25	−1	3	37
3	8	6	15625	1	4	193
15	32768	3	125	−1	4	403
14	16384	1	5	1	3	607
17	131072	7	78125	−1	3	1961
16	65536	5	3125	1	3	2543
8	256	7	78125	1	3	2903
19	524288	2	25	1	4	6473
18	262144	0	1	−1	3	9709
23	8388608	1	5	−1	6	11507
13	8192	8	390625	1	3	14771
22	4194304	8	390625	−1	5	15653
10	1024	11	48828125	1	7	22327
18	262144	9	1953125	1	4	27349
20	1048576	4	625	−1	3	38813
0	1	9	1953125	1	3	72338
21	2097152	6	15625	1	3	78251
5	32	10	9765625	1	3	361691
24	16777216	3	125	1	3	621383
23	8388608	10	9765625	1	3	672379
26	67108864	7	78125	1	4	829469

$p=5,\ p_0=2,\ p_1=3$

x_0	$p_0^{x_0}$	x_1	$p_1^{x_1}$	sign	u	w
12	4096	16	43046721	−1	3	−344341
5	32	15	14348907	−1	3	−114791
7	128	1	3	−1	3	1
6	64	8	6561	1	3	53
14	16384	2	9	−1	3	131
13	8192	9	19683	1	3	223
20	1048576	10	59049	1	3	8861
21	2097152	3	27	−1	3	16777

Table II. (Theorem 6.3.)

			$\mathrm{ord}_p(x)$						$\mathrm{ord}_p(y)$						$\mathrm{ord}_p(z)$					
x	y	z	$p=2$	3	5	7	11	13	$p=2$	3	5	7	11	13	$p=2$	3	5	7	11	13
2401	4160	6561	0	0	0	4	0	0	6	0	1	0	0	1	0	8	0	0	0	0
875	6561	7436	0	0	3	1	0	0	0	8	0	0	0	0	2	0	0	0	1	2
1183	6561	7744	0	0	0	1	0	2	0	8	0	0	0	0	6	0	0	0	2	0
1125	8192	9317	0	2	3	0	0	0	13	0	0	0	0	0	0	0	0	1	3	0
1183	8192	9375	0	0	0	1	0	2	13	0	0	0	0	0	0	1	5	0	0	0
16	14625	14641	4	0	0	0	0	0	0	2	3	0	0	1	0	0	0	0	4	0
81	14560	14641	0	4	0	0	0	0	5	0	1	1	0	1	0	0	0	0	4	0
1936	13689	15625	4	0	0	0	2	0	0	4	0	0	0	2	0	0	6	0	0	0
3718	11907	15625	1	0	0	0	1	2	0	5	0	2	0	0	0	0	6	0	0	0
5824	9801	15625	6	0	0	1	0	1	0	4	0	0	2	0	0	0	6	0	0	0
49	16335	16384	0	0	0	2	0	0	0	3	1	0	2	0	14	0	0	0	0	0
2695	13689	16384	0	0	1	2	1	0	0	4	0	0	0	2	14	0	0	0	0	0
8019	8788	16807	0	6	0	0	1	0	2	0	0	0	0	3	0	0	0	5	0	0
3584	14641	18225	9	0	0	1	0	0	0	0	0	0	4	0	0	6	2	0	0	0
1625	16807	18432	0	0	3	0	0	1	0	0	0	5	0	0	11	2	0	0	0	0
3993	16807	20800	0	1	0	0	3	0	0	0	0	5	0	0	6	0	2	0	0	1
49	28512	28561	0	0	0	2	0	0	5	4	0	0	1	0	0	0	0	0	0	4
12936	15625	28561	3	1	0	2	1	0	0	0	6	0	0	0	0	0	0	0	0	4
22000	6561	28561	4	0	3	0	1	0	0	8	0	0	0	0	0	0	0	0	0	4
15625	17303	32928	0	0	6	0	0	0	0	0	0	0	3	1	5	1	0	3	0	0
507	32768	33275	0	1	0	0	0	2	15	0	0	0	0	0	0	0	2	0	3	0
10985	41503	52488	0	0	1	0	0	3	0	0	0	3	2	0	3	8	0	0	0	0
10000	49049	59049	4	0	4	0	0	0	0	0	0	3	1	1	0	10	0	0	0	0
14641	46875	61516	0	0	0	0	4	0	0	1	6	0	0	0	2	0	0	1	0	3
7168	78125	85293	10	0	0	1	0	0	0	0	7	0	0	0	0	8	0	0	0	1
20449	97200	117649	0	0	0	0	2	2	4	5	2	0	0	0	0	0	0	6	0	0
13	151250	151263	0	0	0	0	0	1	1	0	4	0	2	0	0	2	0	5	0	0
12005	161051	173056	0	0	1	4	0	0	0	0	0	0	5	0	10	0	0	0	0	2
121	255879	256000	0	0	0	0	2	0	0	9	0	0	0	1	11	0	3	0	0	0
2197	583443	585640	0	0	0	0	0	3	0	5	0	4	0	0	3	0	1	0	4	0
91	1771470	1771561	0	0	0	1	0	1	1	11	1	0	0	0	0	0	0	0	6	0

Table III.

$-\log_2 5 / \log_2 3 =$

```
0.10101 11101 00001 11110 11000 10101 00000 01001 11101 00010 10000 10011 10110 10000 01011 11100 00001 11010 00000 00001
  00010 11100 11100 10111 01001 01101 11000 01010 01110 01010 11110 00000 00101 00110 00100 00011 01111 01110 01010 11101
  10010 01001 00001 10100 00111 00001 11111 01111 00011 10110 00000 00101 01101 01100 00010 01110 11100 11101 11011 01011
  10111 00100 11111 00100 00100 10001 10010 01011 10010 01101 00000 01100 01111 10111 01101 00110 11110 00000 01100 11010
  00001 00111 11011 11001 01000 10000 00110 11011 11001 01000 01110 11001 00010 00010 11000 01000 00110 01100 11101 11110
  00000 11101 10101 11100 10010 11101 01011 10001 01100 11000 01100 00110 01001 01100 01000 01001 10001 11000 01101 10100
  00110 00011 0111....
```

$-\log_2 7 / \log_2 3 =$

```
0.01001 01011 01111 11100 11010 01111 11111 10010 01000 11100 00011 00100 01011 11110 00101 11010 11110 10001 00000 01110
  01011 00101 10010 00111 10111 01001 10001 11000 11011 10111 01011 01001 01011 00100 11001 11000 10111 01100 00001 01110
  11000 01011 00011 01110 10000 01101 11000 11001 00010 00011 01100 11110 11010 11110 00001 10110 01010 00001 01000 11011
  10010 01000 00011 01011 10010 11001 10000 01101 10111 01001 01000 00101 00011 11001 10000 00011 00101 00100 00000 10011
  11100 01110 11110 10101 00101 01110 11100 10000 01011 00100 01100 11100 00100 01110 10001 00001 10000 00111 10111 00011
  10111 01100 10110 00111 00101 10011 01100 11111 00101 01011 11101 11011 01011 10110 01100 10001 11011 01100 01001 11111
  10001 11110 1010....
```

$-\log_2 11 / \log_2 3 =$

```
0.10011 01110 00001 01001 00110 01010 01110 00100 00101 10000 01000 11001 10010 01111 10110 00011 00101 01110 00011 11101
  10100 01101 01101 10111 10110 01100 10110 11110 00000 11000 01000 11111 10010 10011 10110 01001 01111 00101 00011 11100
  01100 10100 01000 10101 00010 01011 10111 10000 00001 01000 11010 01010 11010 11100 00001 10110 01001 11111 11110 10010
  11101 11100 10010 00100 11000 00000 01110 10100 00101 10010 10010 10010 01111 10100 01001 10111 01000 10101 01110 01000
  11001 11001 10011 11110 10000 11001 10110 10101 00010 11100 11101 11011 10110 00111 11010 11000 01000 10111 11110 10010
  10101 11100 11011 00111 11000 00111 11100 01101 11100 01101 00001 00100 01001 11000 01101 10101 00110 10101 00001 00001
  10101 11001 1111....
```

$-\log_2 13 / \log_2 3 =$

```
0.11011 10110 10100 10001 01100 01111 10001 00110 01110 00001 01110 01011 11101 10110 11100 00111 11111 11001 11011 11110
  00011 11110 01000 10010 11011 11101 01000 11000 11111 01011 00001 10101 00110 00000 00001 01101 10010 00101 00100 10101
  01011 11100 10011 11011 00000 01110 10100 00011 10000 11101 00010 00000 10011 01011 11011 10000 11010 10010 11010 00100
  10010 00001 01000 10101 10110 01001 01000 11111 01101 00111 11100 10001 01110 11000 01010 10011 11001 01111 01111 00001
  10011 00000 00101 10101 10111 11100 01011 01011 11000 10001 11110 01000 10110 00011 11011 11011 11010 11110 10100 01110
  01000 00111 00000 00000 00011 10111 11111 10100 01111 10010 01100 10101 00111 10110 11010 01111 10000 11101 10111 10001
  00000 01100 1001....
```

Table III. (cont.)

$- \log_3 5 \,/\, \log_3 2 =$

```
0.11022 12121 22001 12010 21102 10210 10022 20212 20010 10112 22201 21021 21022 10000 22020 12012 02022 21001 00012 02020
  21210 12202 12200 00000 10120 00211 12021 10120 02100 10222 22122 01201 21111 11121 11001 20222 10000 20121 22221 01002
  20220 12211 22211 00100 20202 00012 11112 10122 21001 21200 12201 12220 11100 01102 20010 11102 10222 00020 21202 21112
  20201 21100 11212 22222 21120 02020 12121 02122 11111 10001 10220 21022 10012 11212 20001 10211 02120 02122 1....
```

$- \log_3 7 \,/\, \log_3 2 =$

```
0.20101 10202 20011 12121 01102 11100 01210 20120 02122 02012 20202 00121 21200 01201 11120 11211 11212 22100 00100 22201
  20021 11112 01122 00011 22100 00000 22011 11100 12010 22110 12122 00222 10220 21102 20001 02101 00121 11002 11012 12201
  21011 20100 01110 02000 21222 12010 02201 22012 01022 02021 00210 10221 00221 20202 02222 22122 00100 12021 21220 02220
  20000 00002 00111 11221 11002 20102 12212 12012 22122 00211 01210 01102 21010 20121 01020 11111 20002 10122 2....
```

$- \log_3 11 \,/\, \log_3 2 =$

```
0.21112 20101 00222 20222 01212 01100 12100 01201 01111 01212 01210 20121 20001 12021 01122 21202 12020 00212 11102 11002
  01001 10200 22202 02001 20022 10221 00010 10011 22220 01021 02121 00211 22210 21101 22012 11111 02010 00221 00102 20111
  20202 01201 01220 22022 11221 10121 10202 10011 11002 10220 22110 21121 00112 02122 21200 01021 21002 21002 10010 00110
  00101 12202 12000 21012 11010 11020 00222 00012 11201 11010 00122 01120 22200 20112 12122 10202 01211 00210 2....
```

$- \log_3 13 \,/\, \log_3 2 =$

```
0.10221 02211 12122 22010 10002 01221 00121 02020 11201 02021 02112 21010 20122 02001 02112 21012 10222 01002 01200 01211
  10111 21100 12121 11010 02000 02212 11111 21220 22020 02000 01222 12112 02100 10110 20002 10222 02112 20112 11100 00211
  20012 11102 22220 00112 00001 11110 11102 22201 01122 22211 22201 11011 22201 01200 22121 02101 22222 22002 01010 01021
  12020 20111 12102 00011 02002 02000 10211 00222 12202 02202 20212 22012 01222 20220 11211 20021 11111 00000 2....
```

$- \log_5 3 \,/\, \log_5 2 =$

```
0.33002 02003 04411 23120 44012 01011 00044 43204 30340 00023 14333 12413 43420 40302 10202 44104 32433 24432 03021 12311
  34044 40231 04112 33230 00242 14232 14400 31104 42112 44033 11014 44344 12114 44211 32120 43131 34041 00411 34233 41410
  24120 42032 43014 21421 40044 01142 21004 42021 14011 10404 00214 31110 04441 42431 24423 0243....
```

$- \log_5 7 \,/\, \log_5 2 =$

```
0.03044 34433 10114 43203 12033 14002 12341 31312 03421 00343 41423 00040 24241 22103 14240 32214 11401 42230 13040 33404
  04310 43034 13233 23241 43002 44411 41124 22443 42412 30420 11223 43101 01000 42112 10443 34210 03410 14414 02220 24443
  13332 33123 23331 20323 44440 13210 14403 32122 03040 31123 04212 22443 44223 23133 02003 1240....
```

Table III. (cont.)

$-\log_5 11 / \log_5 2 =$

```
0.44032 21012 13124 21134 03320 33422 21041 12112 42420 00220 41143 12040 32144 21100 01304 24013 43401 23313 12022 34404
  12413 10214 30123 11014 24110 42444 42030 02413 20241 22304 23423 13414 03234 30000 10334 44322 00330 01104 44410 44113
  31022 33142 14441 44113 21413 23132 31413 32032 01221 40210 24101 30133 13110 13400 22110 2334....
```

$-\log_5 13 / \log_5 2 =$

```
0.12423 02224 01323 24314 23021 32420 14134 41224 04403 11334 43213 33303 03130 32244 11133 43243 23422 11320 41041 31134
  41320 34110 03024 40012 23213 10014 41441 04420 40114 00021 33224 30103 03243 32031 22021 20234 32441 00013 04203 43134
  22012 30332 22422 43110 13302 34431 13241 13230 44204 14432 33210 24121 13144 03230 14301 3040....
```

$-\log_7 2 / \log_7 3 =$

```
0.20603 14521 11264 52354 45364 60036 13315 13044 46363 40432 02366 04135 21304 53356 32205 44546 66301 00123 63633 04024
  40631 53556 64053 50031 26336 46625 03465 02235 11551 46123 25164 52364 25520 12240 64220 00164 43634 02066 41264 61233
  41326 32413 65633 52502 526....
```

$-\log_7 5 / \log_7 3 =$

```
0.62250 35002 24045 66544 01041 43506 34535 04453 26545 45453 33261 65353 53330 22443 10105 55005 62520 33063 16320 22253
  42306 51054 13301 06465 43020 41555 41121 64255 11350 55053 64515 44465 36222 25605 66346 16142 31340 45522 31033 14255
  21343 24510 62633 00155 361....
```

$-\log_7 11 / \log_7 3 =$

```
0.25035 56505 33553 02331 10224 32143 50543 02561 42352 23430 26326 53446 23462 31210 02416 02335 13066 54240 13006 60451
  00441 56630 66142 56540 44042 52255 15314 10131 40626 10543 02504 04254 45066 50232 65102 41555 41254 02211 54156 34054
  41366 64215 64014 56645 550....
```

$-\log_7 13 / \log_7 3 =$

```
0.21305 11055 56501 22565 55610 32506 13150 66465 56420 46465 21650 16426 11613 41010 66512 26566 50652 02025 40431 56566
  43655 56120 34610 53642 61544 36122 61225 42410 23035 15004 26220 14444 23632 33426 15605 51104 34116 04520 65502 65542
  24255 36603 45452 66563 536....
```

Table III. (cont.)

$-\log_{11} 3 / \log_{11} 2 =$

```
0.08248 A4245 06166 43468 58202 44A56 73171 16758 A203A 8A543 28431 86731 11411 4A296 993A7 31A79 00421 95444 80670 57433
 59439 78064 34745 1A710 64682 08044 81761 27049 03452 3661A 40979 29601 898A4 50....
```

$-\log_{11} 5 / \log_{11} 2 =$

```
0.351A9 7223A 31378 09193 42445 306A3 96588 11862 48667 AA6A2 39A03 77139 01693 21678 33652 12687 95AA8 24190 78276 28711
 08399 68022 2A607 55A17 2231A 80798 76947 73936 21835 30A1A 95324 1A8A3 82999 67....
```

$-\log_{11} 7 / \log_{11} 2 =$

```
0.44804 92167 71327 83472 37453 00781 3256A 2A367 85671 88907 799A1 4AAAA 784A1 29329 A6950 17481 86846 17379 94130 77091
 29354 33161 9A146 03746 52A14 20214 22541 58A91 50337 7954A 89A01 43809 A8152 52....
```

$-\log_{11} 13 / \log_{11} 2 =$

```
0.9011A 94962 52990 39096 3A68A 7556A 1A4A3 44758 57692 20188 42770 072A3 9A977 8819A 97518 14396 07360 899A2 99391 26176
 84077 81181 54473 58532 58A01 91643 28056 63940 99265 27989 37450 85913 91289 56....
```

$-\log_{13} 3 / \log_{13} 2 =$

```
0.621B3 15581 0A077 3B5C8 49202 39A32 82105 848C7 70988 B863B 75151 52114 5C25A 04902 6B6C7 377B9 3122B 5CAC0 13945 A2471
 9B4BA A79C2 7A91C 5A989 C392A CC16A A20A1 75C6B 06BB6 8A3B9 C782B AA70C AB218 C9....
```

$-\log_{13} 5 / \log_{13} 2 =$

```
0.44570 79C51 73665 3796C B7C61 335A0 79906 2B429 51211 4900B 481B1 621AB 2AC77 C2291 1662A BB03A 8CB9C 77331 74992 11C07
 BB101 10301 77310 B8B28 83AB2 57975 7C697 57928 23B72 297CB 0A414 32B3C A67A8 48....
```

$-\log_{13} 7 / \log_{13} 2 =$

```
0.A1C78 9C71A 63110 51424 42CA9 0AAA7 B225B B0281 501B1 976C2 3C05B 09CA3 AB803 C3251 838AC 72502 A1844 03603 644A8 A8501
 173BB BBC1C 30466 223C6 C98B4 564C2 47140 28856 C8676 15C30 12892 A3317 163C8 CA....
```

$-\log_{13} 11 / \log_{13} 2 =$

```
0.1760A A080C 20874 BB876 B2162 75989 CB19B B7CC2 26BB7 87093 5A833 A9375 AB4BA 8C0BC 1A698 96C6B A9411 34B75 4B718 63BC3
 571A9 14566 8819B A4B95 B4244 452A8 29623 49AA5 CB804 AC61A CC513 08855 79185 43....
```

Table IV.

nr.	p	m	#
1	2	44	-
2	3	28	-
3	5	20	-
4	7	16	-
5	11	12	-
6	13	12	-
7	2	33	-
8	3	21	-
9	5	15	-
10	7	12	-
11	11	9	-
12	3	9	-
13	2	22	-
14	3	14	-
15	5	10	-
16	7	8	-
17	11	6	-
18	13	6	-
19	2	21	-
20	2	20	-
21	2	19	-
22	2	18	-
23	2	17	-
24	2	16	-
25	2	15	-
26	2	14	-
27	2	13	1
28	2	12	2
29	2	11	2
30	3	13	-
31	3	12	-
32	3	11	-
33	3	10	1
34	3	9	1
35	3	8	1
36	3	7	6
37	5	9	-
38	5	8	-
39	5	7	-
40	5	6	-
41	5	5	6
42	7	7	-
43	7	6	-
44	7	5	1
45	7	4	4
46	11	5	-
47	11	4	1
48	11	3	4
49	13	5	-
50	13	4	-
51	13	3	1
52	2	10	2
53	2	9	3
54	2	8	6
55	2	7	15
56	2	6	16
57	2	5	26
58	2	4	31
59	2	3	44
60	3	6	5
61	3	5	8
62	3	4	16
63	3	3	35
64	3	2	54
65	3	1	87
66	5	4	1
67	5	3	5
68	5	2	18
69	5	1	36
70	7	3	-
71	7	2	6
72	7	1	18
73	11	2	1
74	11	1	8
75	13	2	-
76	13	1	4

Chapter 7. The sum of two S-units being a square.

7.1. Introduction.

Let $p_1, \ldots, p_s$ ($s \geqslant 1$) be distinct primes, and let S be the set of positive rational integers which have no prime divisors different from the p_i . A rational number is called an *S-unit* if its absolute value is a quotient of elements of S . Thus the set of S-units is

$$\langle \pm p_1^{x_1} \cdot \ldots \cdot p_s^{x_s} \mid x_i \in \mathbb{Z} \text{ for } i = 1, \ldots, s \rangle .$$

We study the diophantine equation

$$x + y = z^2$$

in S-units x, y , and $z \in \mathbb{Q}$, where the set of primes $p_1, \ldots, p_s$ is given. We show how to find all solutions of this equation, using the theory of p-adic linear forms in logarithms, and a computational p-adic diophantine approximation method. We actually perform all the necessary computations for solving the equation completely for $\{ p_1, \ldots, p_s \} = \{ 2, 3, 5, 7 \}$. This type of equations has applications in arithmetic algebraic geometry (cf. Setzer [1975], Pinch [1984]).

We start with getting rid of the denominators. Let x, y, z be a solution. There is a $d \in S$ such that $|d \cdot x|, |d \cdot y| \in S$. Put $d = d_1 \cdot d_2^2$, where $d_1, d_2 \in S$ and d_1 squarefree. Then

$$d_1 \cdot d \cdot x + d_1 \cdot d \cdot y = (d_1 \cdot d_2 \cdot z)^2 ,$$

which has the same form as $x + y = z^2$, but now $|d_1 \cdot d \cdot x|, |d_1 \cdot d \cdot y| \in S \subset \mathbb{Z}$ and $d_1 \cdot d_2 \cdot z \in \mathbb{Z}$. Without loss of generality we may therefore study

$$x + y = z^2 , \tag{7.1}$$

where

$$\begin{cases} x \in S , \quad \pm y \in S , \quad z \in \mathbb{Z} , \\ x \geqslant y , \quad z > 0 , \\ (x, y) \text{ is squarefree} . \end{cases} \tag{7.2}$$

We shall prove the following results.

THEOREM 7.1. *Let* $p_1, \ldots, p_s$ *be given. There exists an effectively computable constant* C *, depending on* $p_1, \ldots, p_s$ *only, such that any solution* x, y, z *of equation* (7.1) *with conditions* (7.2) *satisfies* $\max(x, |y|, z) < C$.

THEOREM 7.2. *Let* $\{ p_1, \ldots, p_s \} = \{ 2, 3, 5, 7 \}$. *Equation* (7.1) *with conditions* (7.2) *has exactly the* 388 *solutions given in* Table I.

Remarks. 1. The Tables are given in Section 7.9. We stress that the aim of this chapter is not only to prove these theorems, but to show as well that for any given set of primes $\{ p_1, \ldots, p_s \}$ a result similar to Theorem 7.2 can be proved along the same lines, in a more or less algorithmic way.
2. Equation (7.1) with conditions (7.2) can be seen as a further generalization of the generalized Ramanujan-Nagell equation

$$x^2 + k = p_1^{n_1} \cdot \ldots \cdot p_s^{n_s} , \tag{7.3}$$

(cf. Chapter 4), namely by taking $|k| \in S$ arbitrary instead of $k \in \mathbb{Z}$ fixed. The method of this chapter to solve (7.1) is also a generalization of the method of Chapter 4 to solve (7.3).

Equation (7.1) can be transformed into a number of Pell-like equations. Put

$$x = D \cdot u^2 ,$$

where $D, u \in S$, and D is squarefree. There are only 2^s possibilities for D . Now, (7.1) is equivalent to a finite number of equations

$$z^2 - D \cdot u^2 = y \tag{7.4}$$

in $u \in S$, $\pm y \in S$, $z \in \mathbb{Z}$, with $z > 0$ and $(u,y) = 1$. We treat equation (7.4) by factorizing its both sides in the field $K = \mathbb{Q}(\sqrt{D})$. When dealing with equation (7.4) we allow z and u to be negative.

7.2. The case $D = 1$.

First we consider the special case $D = 1$. Then (7.4) is equivalent to

$$\begin{cases} z + u = y_1 \\ z - u = y_2 \end{cases},$$

where $y = y_1 \cdot y_2$, $y_1 \in S$, $\pm y_2 \in S$, and $y_1 > |y_2|$. Subtraction yields

$$2 \cdot u = y_1 - y_2 , \tag{7.5}$$

where now all variables u, y_1, y_2 (apart from the sign) are in S , hence in $\mathbb{Z}$. By $(u,y_1) = (u,y_2) = 1$, equation (7.5) is of the form $a + b = c$, or $2 \cdot a + 2 \cdot b = 2 \cdot c$, where a, b, c are composed of primes 2, p_1, ..., p_s only, and $(a,b) = 1$, $a \geqslant b > 0$. In Chapter 6 it was shown how to solve $a + b = c$. For our standard example $\{ p_1, \ldots, p_s \} = \{ 2, 3, 5, 7 \}$ we have the following result.

LEMMA 7.3. *Let* $\{ p_1, \ldots, p_s \} = \{ 2, 3, 5, 7 \}$. *Equation* (7.1) *with conditions* (7.2) *and* $D = 1$ *has exactly the* 95 *solutions given in* Table I *with* $D = 1$.

Proof. From Theorem 6.3 it follows that $a + b = c$ with a, b, $c \in S$, $(a,b) = 1$, $a \geqslant b$ has exactly 63 solutions. They are easy to compute. Each of these gives rise to three possibilities for (7.5):

$$\text{if } 2 \mid a \text{ then } (u,y_1,y_2) = (\tfrac{1}{2}a,b,c),\ (b,2c,2a),\ (c,2a,-2b),$$
$$\text{if } 2 \mid b \text{ then } (u,y_1,y_2) = (a,2b,2c),\ (\tfrac{1}{2}b,c,a),\ (c,2a,-2b),$$
$$\text{if } 2 \mid c \text{ then } (u,y_1,y_2) = (a,2b,2c),\ (b,2c,2a),\ (\tfrac{1}{2}c,a,-b).$$

Of the thus found 189 possibilities, the 95 ones given in Table I with $D = 1$ satisfy $x \geqslant y$ and $z > 0$, whereas the others don't. $\square$

This completes our treatment of the case $D = 1$.

7.3. Towards generalized recurrences.

From now on, let $D > 1$. Put $K = \mathbb{Q}(\sqrt{D})$. Let $\sigma : K \to K$ be the automorphism of K with $\sigma(\sqrt{D}) = -\sqrt{D}$. For any number or ideal X in K we write X' for $\sigma(X)$, for convenience. Let $\mathfrak{p}_i$ for $i = 1, \ldots, s$ be the prime ideal in K such that $\operatorname{ord}_{p_i}(\mathfrak{p}_i) > 0$. If p_i splits in $\mathcal{O}_K$, this is well defined if a choice has been made from the two possibilities for $\sqrt{D} \pmod{p\ }$. Put for a solution z, u, y of (7.4)

$$\chi = z + u\cdot\sqrt{D} .$$

Then $y = \chi\cdot\chi'$, and by $(u,y) = 1$ we have

$$\min\bigl(\operatorname{ord}_{p_i}(u),\ \operatorname{ord}_{p_i}(y)\bigr) = 0 . \tag{7.6}$$

Equation (7.4) leads to the conjugated ideal equations

$$\left\{\begin{aligned} (\chi) &= \prod_{i=1}^{s} \mathfrak{p}_i^{a_i}\cdot\mathfrak{p}_i'^{\,b_i} \\ (\chi') &= \prod_{i=1}^{s} \mathfrak{p}_i'^{\,a_i}\cdot\mathfrak{p}_i^{b_i} \end{aligned}\right. \tag{7.7}$$

where $a_i, b_i \in \mathbb{N}_0$, and $b_i = 0$ if $\mathfrak{p}_i = \mathfrak{p}_i'$. We need the following auxiliary lemma.

<u>LEMMA 7.4.</u> *If* $\xi \in K$ *and* $\operatorname{ord}_p(\xi) = \operatorname{ord}_p(\xi')$ *for a prime* p *, then*

$$\operatorname{ord}_p(\xi) \leqslant \operatorname{ord}_p(\xi-\xi') .$$

Moreover, if $p = 2$ *and* $D \equiv 1 \pmod 8$ *, then*

$$\operatorname{ord}_2(\xi) \leqslant \operatorname{ord}_2((\xi-\xi')/2) ,$$

and, if $p = 2$ *and* $D \equiv 2, 3 \pmod 4$ *, then*

$$\operatorname{ord}_2(\xi) \leqslant \operatorname{ord}_2((\xi-\xi')/2\sqrt{D}) + \tfrac{1}{2} .$$

<u>Proof.</u> This is an easy exercise, which we leave to the reader. □

We distinguish, as usual, three cases for the factorization of the prime p_i in K : it may split, ramify or remain prime. See Borevich and Shafarevich [1966], section III.8.

→ p_i remains prime in K . Then $p_i \nmid D$, and if $p_i = 2$ then $D \equiv 5 \pmod 8$. We have $(p_i) = \mathfrak{p}_i = \mathfrak{p}_i'$, and from $\operatorname{ord}_{p_i}(\chi) = \operatorname{ord}_{p_i}(\chi')$ and Lemma 7.4 we obtain

$$\operatorname{ord}_{p_i}(y) = 2\cdot\operatorname{ord}_{p_i}(\chi) \leqslant 2\cdot\operatorname{ord}_{p_i}(\chi-\chi') = 2\cdot\operatorname{ord}_{p_i}(2\cdot u\cdot\sqrt{D}) .$$

It follows, using (7.6), that

if $p_i \neq 2$ then $\mathrm{ord}_{p_i}(y) = 2\cdot a_i = 0$,

if $p_i = 2$ then $\mathrm{ord}_2(y) = 2\cdot a_i = 0, 2$, and if $a_i = 1$ then $\mathrm{ord}_2(u) = 0$.

$\rightarrow$ p_i ramifies in K . Then $p_i \mid D$ if $p_i \neq 2$, and $D \equiv 2, 3 \pmod 4$ if $p_i = 2$. We have $(p_i) = \mathfrak{p}_i^2$, $\mathfrak{p}_i = \mathfrak{p}_i'$, and $\mathrm{ord}_{p_i}(\chi) = \mathrm{ord}_{p_i}(\chi') = \frac{1}{2}\cdot a_i$. From Lemma 7.4 we find

$$\mathrm{ord}_{p_i}(y) = 2\cdot\mathrm{ord}_{p_i}(\chi) \leqslant 1 + 2\cdot\mathrm{ord}_{p_i}((\chi-\chi')/2\cdot\sqrt{D}) = 1 + 2\cdot\mathrm{ord}_{p_i}(u) .$$

By (7.6) we obtain

$$\mathrm{ord}_{p_i}(y) = a_i = 0, 1 \text{ , and if } a_i = 1 \text{ then } \mathrm{ord}_{p_i}(u) = 0 .$$

$\rightarrow$ p_i splits in K . Then $p_i \nmid D$, and if $p_i = 2$ then $D \equiv 1 \pmod 8$. We have $(p_i) = \mathfrak{p}_i\cdot\mathfrak{p}_i'$, $\mathfrak{p}_i \neq \mathfrak{p}_i'$. Further, $\mathrm{ord}_{p_i}(\mathfrak{p}_i) = 1$, $\mathrm{ord}_{p_i}(\mathfrak{p}_i') = 0$. Hence $\mathrm{ord}_{p_i}(\chi) = a_i$, $\mathrm{ord}_{p_i}(\chi') = b_i$. If $a_i = b_i$ then from

$$\mathrm{ord}_{p_i}(y) = 2\cdot\mathrm{ord}_{p_i}(\chi) \leqslant 2\cdot\mathrm{ord}_{p_i}((\chi-\chi')/2) = 2\cdot\mathrm{ord}_{p_i}(u)$$

we obtain by (7.6) that

$$\mathrm{ord}_{p_i}(y) = a_i = b_i = 0 .$$

If $a_i \neq b_i$ then $\mathrm{ord}_{p_i}(y) = a_i + b_i > 0$, hence $\mathrm{ord}_{p_i}(u) = 0$, by (7.6). We infer in this case

$$\begin{aligned}\mathrm{ord}_{p_i}(y) &= a_i + b_i \geqslant 1 + 2\cdot\min(a_i,b_i) = 1 + 2\cdot\mathrm{ord}_{p_i}(\chi-\chi') \\ &= 1 + 2\cdot\mathrm{ord}_{p_i}(2) .\end{aligned}$$

It follows that

$$\mathrm{ord}_{p_i}(y) = \max(a_i,b_i) \text{ , } \min(a_i,b_i) = 0 \text{ if } p_i \neq 2 ,$$

$$\mathrm{ord}_{p_i}(y) = \max(a_i,b_i) + 1 \text{ , } \min(a_i,b_i) = 1 \text{ if } p_i = 2 .$$

Put $b_0 = \min(a_i,b_i)$ if $p_i = 2$ occurs, and $b_0 = 0$ otherwise. (Note that

$\min(a_i, b_i) = 1$ may occur only if $\mathfrak{p}_i \neq \mathfrak{p}_i'$, hence only if $p_i = 2$ splits). Let us assume that the splitting primes of $p_1, \ldots, p_s$ are $p_1, \ldots, p_t$ for some $0 \leqslant t \leqslant s$. Put

$$I = \{\, i \mid 1 \leqslant i \leqslant t\,,\ a_i > b_i \,\}\,,$$

$$I' = \{\, i \mid 1 \leqslant i \leqslant t\,,\ a_i < b_i \,\}\,.$$

For $i = 1, \ldots, t$, let h_i be the smallest positive integer such that $\mathfrak{p}_i^{h_i}$ is a principal ideal, say

$$\mathfrak{p}_i^{h_i} = (\pi_i)\,.$$

If h denotes the class number of K, then $h_i \mid h$. Now, $\pi_i \in K$ is determined up to multiplication by a unit. Thus we may choose π_i such that

$$|\pi_i| > |\pi_i'| \quad \text{if} \quad i \in I\,,$$

$$|\pi_i| < |\pi_i'| \quad \text{if} \quad i \in I'\,.$$

For $i = 1, \ldots, t$, put

$$|\, a_i - b_i \,| = c_i \cdot h_i + d_i\,,$$

with $c_i, d_i \in \mathbb{N}_0$, and $0 \leqslant d_i \leqslant h_i - 1$. Consider the ideal

$$\mathfrak{a} = (2)^{b_0} \cdot \prod_{i \in I} \mathfrak{p}_i^{d_i} \cdot \prod_{i \in I'} \mathfrak{p}_i'^{\,d_i} \cdot \prod_{i=t+1}^{s} \mathfrak{p}_i^{a_i}\,.$$

From the above considerations it follows that, for given K, $p_1, \ldots, p_s$, there are only finitely many possibilities for $\mathfrak{a}$. By (7.7) it follows that

$$(\chi) = \mathfrak{a} \cdot \prod_{i \in I} (\pi_i)^{c_i} \cdot \prod_{i \in I'} (\pi_i')^{c_i}$$

(namely, $|a_i - b_i| = \max(a_i, b_i)$ if $p_i \neq 2$, since then $\min(a_i, b_i) = 0$; and $|a_i - b_i| = \max(a_i, b_i) - 1$ if $p_i = 2$ and $b_0 = 1$). Hence $\mathfrak{a}$ is a principal ideal, say

$$\mathfrak{a} = (\alpha)$$

for an $\alpha \in \mathcal{O}_K$. Up to multiplication by a unit, there are only finitely many possibilities for α. Let ε be the fundamental unit of K with $\varepsilon > 1$.

Now, (7.7) leads to the system of equations

$$\left\{\begin{array}{l} \chi \;= z + u\sqrt{D} = \pm\alpha \cdot \varepsilon^{n} \cdot \prod_{i\in I} \pi_i^{c_i} \cdot \prod_{i\in I'} \pi_i'^{\,c_i} \\ \\ \chi' = z - u\sqrt{D} = \pm\alpha' \cdot \varepsilon'^{\,n} \cdot \prod_{i\in I} \pi_i'^{\,c_i} \cdot \prod_{i\in I'} \pi_i^{c_i} \end{array}\right. , \tag{7.8}$$

where $n \in \mathbb{Z}$. Put for $n \in \mathbb{Z}$, $m_1, \ldots, m_t \in \mathbb{N}_0$, and for each possible α

$$G_\alpha(n, m_1, \ldots, m_t) = \frac{\alpha}{2\sqrt{D}} \cdot \varepsilon^{n} \cdot \prod_{i\in I} \pi_i^{m_i} \cdot \prod_{i\in I'} \pi_i'^{\,m_i} - \frac{\alpha'}{2\sqrt{D}} \cdot \varepsilon'^{\,n} \cdot \prod_{i\in I} \pi_i'^{\,m_i} \cdot \prod_{i\in I'} \pi_i^{m_i},$$

$$H_\alpha(n, m_1, \ldots, m_t) = \frac{\alpha}{2} \cdot \varepsilon^{n} \cdot \prod_{i\in I} \pi_i^{m_i} \cdot \prod_{i\in I'} \pi_i'^{\,m_i} + \frac{\alpha'}{2} \cdot \varepsilon'^{\,n} \cdot \prod_{i\in I} \pi_i'^{\,m_i} \cdot \prod_{i\in I'} \pi_i^{m_i}.$$

Then (7.8) is equivalent to

$$\left\{\begin{array}{l} \pm\, u = G_\alpha(n, c_1, \ldots, c_t) \\ \pm\, z = H_\alpha(n, c_1, \ldots, c_t) \end{array}\right. . \tag{7.9}$$

The functions G_α and H_α are generalized recurrences in the sense that if all variables but one are fixed, then they become integral binary recurrence sequences. We show an example in Fig. 8.

***********	***********	***********	***********	***********	***********	***********	***********	***********	***********	***********
51622180721.	***********	***********	***********	***********	***********	***********	***********	***********	***********	***********
2351330961.	4809762343.	9838603829.	20125344747.	41167375801.	84209878223.	***********	***********	***********	***********	***********
107100421.	219079143.	448137089.	916686307.	1875126621.	3835663103.	7846036249.	16049450427.	32829934901.	67155236503.	***********
4878301.	9978803.	20412129.	41754007.	85409861.	174710043.	357377929.	731034047.	1495366701.	3058865923.	6257545649.
222201.	454523.	929749.	1901847.	3890321.	7957843.	16278189.	33298607.	68132521.	139813803.	296665349.
10121.	20703.	42349.	86627.	177201.	362503.	742229.	1535307.	3548761.	17037743.	269092029.
461.	943.	1929.	3947.	8101.	17223.	50849.	478147.	9940221.	235016543.	5623359289.
21.	43.	89.	207.	1021.	16403.	376449.	8983927.	215136101.	5153326203.	***********
1.	3.	29.	607.	14361.	343643.	8231029.	197168247.	4723054001.	***********	***********
1.	23.	549.	13147.	314921.	7543743.	180706189.	4328717507.	***********	***********	***********
21.	503.	12049.	288627.	6913901.	165618703.	3967305129.	95034616907.	***********	***********	***********
461.	11043.	264529.	6336647.	151790901.	3636067723.	87100006649.	***********	***********	***********	***********
10121.	242443.	5807589.	139117607.	3332485921.	79827871203.	***********	***********	***********	***********	***********
222201.	5322703.	127502429.	3054250707.	73162899361.	***********	***********	***********	***********	***********	***********
4878301.	116857023.	2799245849.	67054397947.	***********	***********	***********	***********	***********	***********	***********
107100421.	2565531803.	61455906249.	***********	***********	***********	***********	***********	***********	***********	***********
2351330961.	56324842643.	***********	***********	***********	***********	***********	***********	***********	***********	***********
51622180721.	***********	***********	***********	***********	***********	***********	***********	***********	***********	***********
***********	***********	***********	***********	***********	***********	***********	***********	***********	***********	***********
***********	***********	***********	***********	***********	***********	***********	***********	***********	***********	***********

Figure 8. $G_\alpha(n,m) = \frac{\alpha}{2\sqrt{D}} \cdot \varepsilon^{n} \cdot \pi^{m} - \frac{\alpha'}{2\sqrt{D}} \cdot \varepsilon'^{\,n} \cdot \pi'^{\,m}$ for $D = 30$, $\alpha = 5 + \sqrt{30}$, $\varepsilon = 11 + 2\cdot\sqrt{30}$, $\pi = 13 + 2\cdot\sqrt{30}$, with $-10 \leqslant n \leqslant 10$ (vertically) and $0 \leqslant m \leqslant 10$ (horizontally). Numbers $\geqslant 10^{12}$ are denoted by asterisks.

7.4. Towards linear forms in logarithms.

Let us write $u_i = \mathrm{ord}_{p_i}(u)$ for $i = 1, \ldots, s$. Put for each α

$$I_U = \langle\, i \mid 1 \leqslant i \leqslant s \ ,\ \mathrm{ord}_{p_i}(G_\alpha(n,m_1,\dots,m_t)) > 0 \ \text{ occurs}$$
$$\text{for at least one } (n,m_1,\dots,m_t) \in \mathbb{Z} \times \mathbb{N}_0^t \,\rangle \ .$$

Note that since $(u,y) = 1$ the sets I_U, I, I' are disjunct. We proceed with the first equation of system (7.9). Written out in full detail it reads

$$\frac{\alpha}{2\sqrt{D}}\cdot\varepsilon^n\cdot\prod_{i\in I}\pi_i^{c_i}\cdot\prod_{i\in I'}\pi_i'^{\,c_i} - \frac{\alpha'}{2\sqrt{D}}\cdot\varepsilon'^{\,n}\cdot\prod_{i\in I}\pi_i'^{\,c_i}\cdot\prod_{i\in I'}\pi_i^{c_i} = \pm\prod_{i\in I_U} p_i^{u_i} \ . \qquad (7.10)$$

Now, I, I', I_U depend on α , which depends on the particular solution of equation (7.4) that we presupposed. However, we know that α belongs to a finite set, which can be computed explicitly. So if we can solve (7.10) completely for each α of this set, then we can find all solutions of (7.9), hence of (7.1).

The set of the α's may be reduced, without loss of generality, as follows. If $D \equiv 1 \pmod 8$ then $b_0 = 0,\ 1$ may both occur, with $\alpha = \alpha_0,\ 2\cdot\alpha_0$ respectively. We only have to consider $2\cdot\alpha_0$, because if $u = u_0,\ z = z_0$ is a solution of (7.9) for $\alpha = \alpha_0$, then $u = 2\cdot u_0,\ z = 2\cdot z_0$ is a solution of (7.9) for $\alpha = 2\cdot\alpha_0$. Hence it is not necessary to consider $\alpha = \alpha_0$ if also $\alpha = 2\cdot\alpha_0$ is already being considered. By the same argument, if $D \equiv 5 \pmod 8$ then with $\alpha = \alpha_0$ such that $\mathrm{ord}_2(\alpha_0) = 0$ also $\alpha = 2\cdot\alpha_0$ may occur, so that we only have to consider the latter. Note that it may now occur that $(u,y) = 2$. The condition $(u,y) = 1$ is used only to ensure that I_U and $I \cup I'$ are disjunct. This remains true in the above cases with $(u,y) = 2$. Further, if $(\alpha_0) \neq (\alpha_0')$ for some α_0 , then we only have to consider one α of the pair $\alpha_0,\ \alpha_0'$. Namely, if the I, I' belonging to α_0 are I_0, I_0' , then the I, I' belonging to α_0' are I_0', I_0 , and then

$$\begin{aligned} G_{\alpha_0'}(n,m_1,\dots,m_t) &= \frac{\alpha_0'}{2\sqrt{D}}\cdot\varepsilon^n\cdot\prod_{I_0'}\pi_i^{c_i}\cdot\prod_{I_0}\pi_i'^{\,c_i} - \frac{\alpha_0}{2\sqrt{D}}\cdot\varepsilon'^{\,n}\cdot\prod_{I_0'}\pi_i'^{\,c_i}\cdot\prod_{I_0}\pi_i^{c_i} \\ &= \pm\left(\frac{\alpha_0'}{2\sqrt{D}}\cdot\varepsilon'^{\,-n}\cdot\prod_{I_0}\pi_i'^{\,c_i}\cdot\prod_{I_0'}\pi_i^{c_i} - \frac{\alpha_0}{2\sqrt{D}}\cdot\varepsilon^{-n}\cdot\prod_{I_0}\pi_i^{c_i}\cdot\prod_{I_0'}\pi_i'^{\,c_i}\right) \\ &= \mp\, G_{\alpha_0}(-n,m_1,\dots,m_t)\ , \end{aligned}$$

(by using $\varepsilon\cdot\varepsilon' = \pm 1$), and analogously

$$H_{\alpha_0'}(n,m_1,\dots,m_t) = \pm\, H_{\alpha_0}(-n,m_1,\dots,m_t)\ .$$

From equation (7.10) we now derive p_i-adic linear forms in logarithms, in three different ways, according to $i \in I, I'$ or I_U. Put

$$\gamma_i = \tfrac{3}{2} \text{ if } p_i = 2, \quad \gamma_i = 1 \text{ if } p_i = 3, \quad \gamma_i = \tfrac{1}{2} \text{ if } p_i \geqslant 5 .$$

Then $\gamma_i > 1/(p_i-1)$, hence if $\mathrm{ord}_{p_i}(\xi) \geqslant \gamma_i$ for a $\xi \in K$ then

$$\mathrm{ord}_{p_i}(\log_{p_i}(1\pm\xi)) = \mathrm{ord}_{p_i}(\xi) . \tag{7.11}$$

We now have the following result.

<u>LEMMA 7.5.</u> *Let* n, c_i $(i \in I \cup I')$, u_i $(i \in I_U)$ *satisfy* (7.10).
(i). *For* $i \in I_U$ *put*

$$\lambda_i = \mathrm{ord}_{p_i}(2\sqrt{D}/\alpha') ,$$

$$\Lambda_i = \log_{p_i}\left(\frac{\alpha}{\alpha'}\right) + n\cdot\log_{p_i}\left(\frac{\varepsilon}{\varepsilon'}\right) + \sum_{j\in I} c_j\cdot\log_{p_i}\left(\frac{\pi_j}{\pi'_j}\right)$$

$$- \sum_{j\in I'} c_j\cdot\log_{p_i}\left(\frac{\pi_j}{\pi'_j}\right) .$$

If $u_i + \lambda_i \geqslant \gamma_i$ *then*

$$u_i + \lambda_i = \mathrm{ord}_{p_i}(\Lambda_i) .$$

(ii). *For* $i \in I$ *put*

$$\kappa_i = \mathrm{ord}_{p_i}\left(\frac{\alpha}{\alpha'}\right) ,$$

$$K_i = \log_{p_i}\left(\frac{\alpha'}{2\sqrt{D}}\right) + n\cdot\log_{p_i}(\varepsilon') - \sum_{j\in I_U} u_j\cdot\log_{p_i}(p_j)$$

$$+ \sum_{j\in I} c_j\cdot\log_{p_i}(\pi'_j) + \sum_{j\in I'} c_j\cdot\log_{p_i}(\pi_j) .$$

If $h_i\cdot c_i + \kappa_i \geqslant \gamma_i$ *then*

$$h_i\cdot c_i + \kappa_i = \mathrm{ord}_{p_i}(K_i) .$$

(ii'). *For* $i \in I'$ *put*

$$\kappa'_i = \mathrm{ord}_{p_i}\left(\frac{\alpha'}{\alpha}\right) ,$$

$$K'_i = \log_{p_i}(\frac{\alpha}{2\sqrt{D}}) + n\cdot\log_{p_i}(\varepsilon) - \sum_{j\in I_U} u_j\cdot\log_{p_i}(p_j)$$

$$+ \sum_{j\in I} c_j\cdot\log_{p_i}(\pi_j) + \sum_{j\in I'} c_j\cdot\log_{p_i}(\pi'_j) \ .$$

If $h_i\cdot c_i + \kappa'_i \geqslant \gamma_i$ *then*

$$h_i\cdot c_i + \kappa'_i = \mathrm{ord}_{p_i}(K'_i) \ .$$

Remark. Note that all the above p_i-adic logarithms are well-defined, since their arguments have p_i-adic order zero. This follows from the fact that I_U, I and I' are disjunct, and if $D \equiv 1 \pmod 8$ from the choice $\alpha = 2\cdot\alpha_0$.

Proof. For (i), divide (7.10) by its second term. For (ii), divide (7.10) by its second term, and add 1. For (ii'), divide (7.10) by its first term, and add -1. Then in all three cases take the p_i-adic order, and apply (7.11). □

The linear forms in logarithms Λ_i, K_i, K'_i , as they appear in Lemma 7.5, seem to be inhomogeneous, since the first term has coefficient 1. However, it can be made homogeneous by incorporating this first term in the other ones, as follows. Put

$$h^* = \mathrm{lcm}\ (\ 2,\ h_1,\ \dots,\ h_s\ ) \ .$$

Note that, by the definition of α ,

$$\alpha^{h^*} = \pm\ \varepsilon^{n_0}\cdot\prod_{i\in I}\pi_i^{n_i}\cdot\prod_{i\in I'}\pi_i'^{n_i}\cdot\prod_{i=t+1}^{s}p_i^{n_i}\cdot 2^{h^*\cdot b_0} \ , \tag{7.12}$$

where the exponents n_i for $0 \leqslant i \leqslant s$ are integral. It follows that

$$\left(\frac{\alpha}{\alpha'}\right)^{h^*} = \pm\left(\frac{\varepsilon}{\varepsilon'}\right)^{n_0}\cdot\prod_{i\in I}\left(\frac{\pi}{\pi'}\right)^{n_i}\cdot\prod_{i\in I'}\left(\frac{\pi'}{\pi}\right)^{n_i} \ .$$

Put

$$\Lambda_i^* = h^*\cdot\Lambda_i \ , \quad n^* = h^*\cdot n + n_0 \ , \quad c_j^* = h^*\cdot c_j + n_j \ .$$

Then it follows that

$$\Lambda_i^* = n^*\cdot\log_{p_i}(\frac{\varepsilon}{\varepsilon'}) + \sum_{j\in I} c_j^*\cdot\log_{p_i}(\frac{\pi_j}{\pi'_j}) - \sum_{j\in I'} c_j^*\cdot\log_{p_i}(\frac{\pi_j}{\pi'_j}) \ .$$

Note that the prime divisors of D are just the ramifying primes. By (7.12),

$$\left(\frac{\alpha}{2\sqrt{D}}\right)^{h^*} = \pm\, \varepsilon^{n_0} \cdot \prod_{i\in I} \pi_i^{n_i} \cdot \prod_{i\in I'} \pi_i'^{\,n_i} \cdot \prod_{i=t+1}^{s} p_i^{n_i-\nu_i} \cdot 2^{h^*\cdot(b_0-\nu_0)} ,$$

where $\nu_i = \frac{1}{2}\cdot h^*\cdot \operatorname{ord}_{p_i}(4D) \in \mathbb{Z}$ for $i = t+1, \ldots, s$, and $\nu_0 = 1$ if 2 splits, $\nu_0 = 0$ otherwise. If $p_i = 2$ splits we have assumed that $b_0 = 1$. Hence the last factor vanishes. So put

$$K_i^* = h^*\cdot K_i \;, \quad K_i'^* = h^*\cdot K_i' \;, \quad u_j^* = h^*\cdot u_j - (\, n_j - \nu_j\,) \;,$$

$$I_U^* = I_U \cup \langle\, i \mid t+1 \leqslant i \leqslant s \;,\; \nu_i \neq 0 \,\rangle \;.$$

Then it follows that

$$K_i^* = n^*\cdot \log_{p_i}(\varepsilon') - \sum_{j\in I_U^*} u_j^*\cdot \log_{p_i}(p_j) + \sum_{j\in I} c_j^*\cdot \log_{p_i}(\pi_j') + \sum_{j\in I'} c_j^*\cdot \log_{p_i}(\pi_j) \;,$$

$$K_i'^* = n^*\cdot \log_{p_i}(\varepsilon) - \sum_{j\in I_U^*} u_j^*\cdot \log_{p_i}(p_j) + \sum_{j\in I} c_j^*\cdot \log_{p_i}(\pi_j) + \sum_{j\in I'} c_j^*\cdot \log_{p_i}(\pi_j') \;.$$

This leads to the following reformulation of Lemma 7.5.

<u>LEMMA 7.6.</u> *Let* n, c_i *for* $i \in I \cup I'$, u_i *for* $i \in I_U$ *be a solution of* (7.10), *let* λ_i, κ_i, κ_i' *be as in Lemma* 7.5, *and let* h^*, Λ_i^*, K_i^*, $K_i'^*$, n_i^*, c_i^*, u_i^*, I_U^* *be as above.*

(i). *Let* $i \in I_U$. *If* $u_i + \lambda_i \geqslant \gamma_i$ *then*

$$u_i + \lambda_i + \operatorname{ord}_{p_i}(h^*) = \operatorname{ord}_{p_i}(\Lambda_i^*) \;.$$

(ii). *Let* $i \in I$. *If* $h_i\cdot c_i + \kappa_i \geqslant \gamma_i$ *then*

$$h_i\cdot c_i + \kappa_i + \operatorname{ord}_{p_i}(h^*) = \operatorname{ord}_{p_i}(K_i^*) \;.$$

(ii'). *Let* $i \in I'$. *If* $h_i\cdot c_i + \kappa_i' \geqslant \gamma_i$ *then*

$$h_i\cdot c_i + \kappa_i' + \operatorname{ord}_{p_i}(h^*) = \operatorname{ord}_{p_i}(K_i'^*) \;.$$

Remark. We will study the linear forms in logarithms Λ_i^*, K_i^*, $K_i'^*$ for arbitrary integral values of the variables n^*, c_i^*, u_i^*. Notice that the parameter α has disappeared completely from these linear forms. This means that we have to consider the linear forms for each D only, instead of for each α.

7.5. Upper bounds for the solutions: outline.

Let us first give a global explanation of our application of the theory of p-adic linear forms in logarithms, that gives explicit upper bounds for the variables occurring in the linear forms Λ_i^*, K_i^*, $K_i'^*$. Then we give arguments why we choose this way to apply the theory, and not other possible ways. In the next section we give full details of the derivation of the upper bounds. In the sequel, by the 'constants' $C_1, \ldots, C_{12}$ we mean numbers that depend only on the parameters of (7.10), not on the unknowns n, c_i, u_i.

Put

$$M = \max_{i \in I \cup I'} (c_i) , \quad U = \max_{i \in I_U} (u_i) , \quad B = \max (M, U, |n|) ,$$

$$M^* = \max_{i \in I \cup I'} (c_i^*) , \quad U^* = \max_{i \in I_U} (u_i^*) , \quad B^* = \max (M^*, U^*, |n^*|) ,$$

$$N = \max (|n_0|, \ldots, |n_t|, |n_{t+1}-\nu_{t+1}|, \ldots, |n_s-\nu_s|) .$$

Then it follows that

$$X^* \leqslant h^* \cdot X + N , \quad X \leqslant \frac{X^* + N}{h^*} \tag{7.13}$$

for $X = M, U, B$. We apply Lemma 2.6 to the p-adic linear forms in logarithms. For Λ_i^* we find, in view of Lemma 7.6(i),

$$U < C_1 + C_2 \cdot \log(B^*) , \tag{7.14}$$

and for K_i^*, $K_i'^*$ we find, in view of Lemma 7.6(ii),(ii'),

$$M < C_3 + C_4 \cdot \log(B^*) . \tag{7.15}$$

Here, C_1, C_2, C_3, C_4 are constants that can be written down explicitly. In order to find an upper bound for B we try to find C_{10}, C_{11} such that

$$B < C_{10} + C_{11}\cdot\log(B^{*}) \ . \tag{7.16}$$

In view of (7.13) we may insert and delete asterisks any time we like, as long as we don't specify the constants. In order to prove (7.16) it remains, in view of (7.14) and (7.15), to bound $|n|$ by a constant times $\log B$. We will introduce certain constants C_5, C_6, C_7 , and distinguish three cases:

$$\begin{aligned} &(a). \quad -(C_6 + C_7\cdot M) \leqslant n \leqslant C_5 \ , \\ &(b). \quad n > C_5 \ , \\ &(c). \quad n < -(C_6 + C_7\cdot M) \ . \end{aligned} \tag{7.17}$$

In case (a) it is, by (7.15), obvious that (7.16) holds. In cases (b) and (c) one of the two terms of G_α dominates. We shall show that there exist constants C_8, C_9 such that

$$|n| < C_8 + C_9\cdot U \ . \tag{7.18}$$

Then (7.16) follows from (7.14).

From (7.16) we derive immediately an explicit upper bound C_{12} for B , hence for all the variables involved. Since the constants $C_1, \ldots, C_4$ will be very large, also C_{12} will be very large. To find all solutions we proceed by reducing this upper bound, by applying the computational p-adic diophantine approximation technique described in Section 3.11, to the p-adic linear forms in logarithms Λ_i^*, K_i^*, $K_i'^*$. Crucial in that line of argument is that the constants $C_5, \ldots, C_9$ are very small compared to $C_1, \ldots, C_4$. This method leads to reduced bounds for the p-adic orders of the linear forms. Then we can replace (7.14) and (7.15) by much sharper inequalities, and repeat the above argument, to find a much sharper inequality for (7.16). In general we expect that it is in this way possible to reduce in one step the upper bound C_{12} for B to a reduced bound of size $\log C_{12}$.

Before going into detail we explain briefly that it is possible to treat (7.10) partly by the theory of real (instead of p-adic) linear forms in logarithms, and subsequently by a real computational diophantine approximation technique (cf. Section 3.7), and why we prefer not to do so. First, note that K_i and K_i' have generically more terms than Λ_i , and are therefore more complicated to handle. Since K_i, K_i' occur only in case (a), this is the most difficult case. Equation (7.10) consist of three terms, each of which is purely exponential, i.e. the bases are fixed and the exponents are variable. If one of these three terms is essentially smaller than the other two (more specifically, smaller than the other terms raised to the power δ, for a fixed $\delta \in (0,1)$), then we can apply the real method. There are two ways of doing this. Write (7.10) as

$$\chi - \chi' = 2 \cdot u \cdot \sqrt{D} \ .$$

First, suppose that $|\chi - \chi'| < |\chi'|^{\delta}$. Then $|n|$ cannot be very large, and we are essentially (i.e. apart from a finite domain) in case (a). Unfortunately, the region for $|n|$ that we can cover in this way becomes smaller as $M \to \infty$ (see the example below). Second, suppose that $|\chi| > |\chi'|^{1/\delta}$, or $|\chi| < |\chi'|^{\delta}$. Then we are essentially in case (b) or (c). But this area can be dealt with easier p-adically, since here we use the linear forms Λ_i, whereas the real linear forms in logarithms used in this case will generically have more terms. The areas sketched above, in which we can apply the real theory, will not cover the whole domain corresponding to case (a) (cf. the white regions in Fig. 9 below). Hence we cannot avoid working with the p-adic linear forms K_i, K'_i. But then it is more convenient to avoid the use of real linear forms.

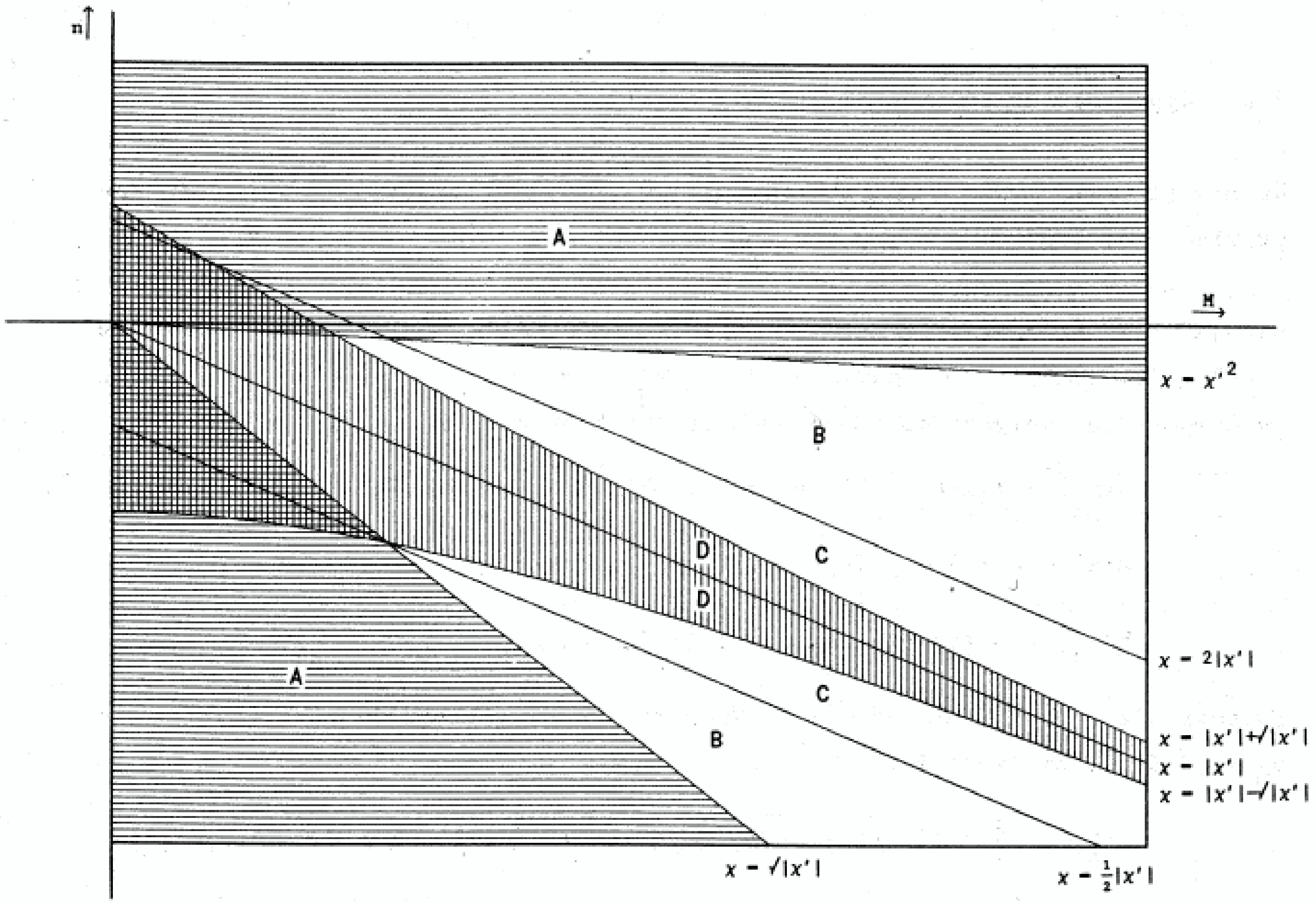


Figure 9.

Let us illustrate the above reasoning with an example. Let $\alpha = \alpha' = 1$, $\varepsilon = 1+\sqrt{2}$, $\pi_1 = 1+2\cdot\sqrt{2}$, $s = 1$, $I = \{1\}$, $p_1 = 7$, $I' = \emptyset$, and $\delta = \frac{1}{2}$. Then we have $\chi = (1+\sqrt{2})^n\cdot(1+2\cdot\sqrt{2})^M$. Fig. 9 above gives in the (n,M)-plane the curves $\chi = \chi'^2,\ 2\cdot|\chi'|,\ |\chi'|+\sqrt{|\chi'|},\ |\chi'|,\ |\chi'|-\sqrt{|\chi'|},\ \frac{1}{2}\cdot|\chi'|,\ \sqrt{|\chi'|}$, which are boundaries of the four regions A, B, C, D. We have the following possibilities.

region	case (essentially)	number of terms in linear form: p-adic method	number of terms in linear form: real method
A	(b),(c)	2	3
B	(b),(c)	2	-
C	(a)	3	-
D	(a)	3	2

The hardest part is C. It can be reduced to $\frac{1}{c}\cdot|\chi'| < \chi < |\chi'| - |\chi'|^\delta$ and $|\chi'| + |\chi'|^\delta < \chi < c\cdot|\chi'|$ for any $c > 1$, $\delta \in (0,1)$, but will never disappear. So we cannot avoid the p-adic linear form in case (a), which then works in regions C and D together.

7.6. Upper bounds for the solutions: details.

We now proceed with filling in the details of the procedure outlined in the previous section.

We apply Yu's lemma (Lemma 2.6) as follows. We have $L = K = \mathbb{Q}(\sqrt{D})$, so $d = 2$. For the α_i we have ε/ε', π_j/π'_j, or ε, ε', p_j, π_j, π'_j. We have to compute the heights of these numbers. We have at once

$$h(p_j) = \log(p_j) \quad \text{if} \quad p_j \geq 3, \quad h(2) = 1,$$

$$h(\varepsilon) = h(\varepsilon') = \tfrac{1}{2}\cdot\log(\varepsilon),$$

$$h(\pi_j) = h(\pi'_j) = \tfrac{1}{2}\cdot\log\bigl(\max(1,|\pi_j|)\cdot\max(1,|\pi'_j|)\bigr).$$

Further, let $\beta = \varepsilon$ or $\beta = \pi_j$. Then the leading coefficient of β/β' is $a_0 = |\beta\cdot\beta'|$, and we infer

$$h\left(\frac{\beta}{\beta'}\right) = \tfrac{1}{2}\log\left(|\beta\cdot\beta'|\cdot\max\left(1,\left|\frac{\beta}{\beta'}\right|\right)\cdot\max\left(1,\left|\frac{\beta'}{\beta}\right|\right)\right) = \log\bigl(\max(|\beta|,|\beta'|)\bigr).$$

Hence

$$h(\frac{\varepsilon}{\varepsilon'}) = \log(\varepsilon) \ , \quad h(\frac{\pi_j}{\pi'_j}) = \log\big(\max(|\pi_j|, |\pi'_j|)\big) \ .$$

The order of the α_i is important in two respects: it is required that the V_i for $i = 1, \ldots, n-1$ are in increasing order, and that $\mathrm{ord}_p(b_n)$ is minimal among the $\mathrm{ord}_p(b_i)$. Since the b_i are the unknowns, we should assume that $V_n \leqslant V_1 \leqslant \ldots \leqslant V_{n-1}$. In the final bound however, only the product $V_1 \cdot \ldots \cdot V_n$ and V^+_{n-1} appear. So the ordering of the V_i only matters for defining V^+_{n-1} . It follows that we can take

$$V_i = \max\ \big(\ h(\alpha_i),\ f_{\mathfrak{p}} \cdot (\log p)/d\ \big)\ ,$$

with the α_i in any order, if we define

$$V^+_{n-1} = \max\ (\ 1,\ V_1,\ \ldots,\ V_n\ )\ .$$

Further, we take

$$B = B_0 = B_n = B' = \max\ \big(\ |b_1|,\ \ldots,\ |b_n|,\ 2,\ \tfrac{4}{3} \cdot n \cdot (p^{f_{\mathfrak{p}}/d} - 1)\ \big)\ .$$

Then $\log(1+\frac{3}{4n} \cdot B) \geqslant f_{\mathfrak{p}} \cdot (\log p)/d$. By $B \geqslant 2$ it follows that $1 + \frac{3}{4n} \cdot B < B$. Hence we can take

$$W = \log B\ .$$

There are two more conditions to be checked. The first one is that $\alpha_1^{b_1} \cdot \ldots \cdot \alpha_n^{b_n} \neq 1$. This is immediate, if we assume the obvious condition that not all b_i are zero. The second one is $[K(\alpha_1^{1/q}, \ldots, \alpha_n^{1/q}):K] = q^n$, which is less obvious. For our situation it follows from the following lemma. Application of Yu's newest results avoids such a condition (cf. Yu [1989]). Nevertheless we include the lemma here, to show that it is often possible to prove such a condition, which may in some cases lead to lower constants.

<u>LEMMA 7.7.</u> *Let* $K = \mathbb{Q}(\sqrt{D})$ *, with* ε *as fundamental unit, and* h *as class number. Let* $p_1, \ldots, p_s$ *be distinct prime numbers, and let* $\mathfrak{p}_i$ *be for* $i = 1, \ldots, s$ *a prime ideal in* K *lying above* p_i *. Let* h_i *be a divisor of* h *such that* $\mathfrak{p}_i^{h_i}$ *is principal, and denote a generator by* π_i *. Let either*: (1) *all* p_i *split, and then*

$$\xi_0 = \frac{\varepsilon}{\varepsilon'}\ , \quad \xi_j = \frac{\pi_j}{\pi'_j} \quad \text{for} \quad i = 1, \ldots, s\ ,$$

or: (2)

$$\xi_0 = \varepsilon \text{ or } \varepsilon' , \quad \xi_j = \pi_j \text{ or } \pi'_j \text{ for } j = 1, \ldots, s .$$

Let q *be an odd prime, not dividing* h *. Then*

$$[K(\xi_0^{1/q}, \ldots, \xi_s^{1/q}):K] = q^{s+1} .$$

Proof. Let $K_0 = K(\xi_0^{1/q})$, and $K_i = K_{i-1}(\xi_i^{1/q})$ for $i = 1, \ldots, s$. We use induction on i to prove that $[K_s:K] = q^{s+1}$. Note that $[K_0:K] = q$. Suppose that $[K_i:K] = q^{i+1}$. It remains to prove that $[K_{i+1}:K_i] = q$, hence it suffices to prove that $\xi_{i+1} \notin K_i$, since q is prime. Suppose the contrary is true. K_i is a K-vector space of dimension q^{i+1} , with as basis all the elements

$$\tau_{k_0,\ldots,k_i} = \prod_{j=0}^{i} \xi_j^{k_j/q}$$

for $k_j \in \{ 0, 1, \ldots, q-1 \}$ for $j = 0, \ldots, i$. It follows that there exist $a_{k_0,\ldots,k_i} \in K$ such that

$$\xi_{i+1}^{1/q} = \sum_{k_0,\ldots,k_i} a_{k_0,\ldots,k_i} \cdot \tau_{k_0,\ldots,k_i} . \tag{7.19}$$

The group of K-embeddings of K_i into $\mathbb{C}$ is generated by the σ_j for $j = 0, \ldots, i$ defined by

$$\sigma_j(\xi_\ell^{1/q}) = \xi_\ell^{1/q} \quad \text{for } \ell = 0, \ldots, i , \quad \ell \neq j ,$$

$$\sigma_j(\xi_j^{1/q}) = \rho \cdot \xi_j^{1/q} ,$$

where ρ is a primitive q th root of unity. Hence all the embeddings are given by

$$\varphi_{\ell_0,\ldots\ell_i} = \prod_{j=0}^{i} \sigma_j^{\ell_j}$$

for $\ell_j \in \{ 0, 1, \ldots, q-1 \}$. It follows that

$$\varphi_{\ell_0,\ldots,\ell_i}(\tau_{k_0,\ldots,k_i}) = \prod_{j=0}^{i} \sigma_j^{\ell_j}\Big(\prod_{m=0}^{i} \xi_m^{k_m/q} \Big) = \prod_{j=0}^{i} \rho^{\ell_j k_j} \cdot \tau_{k_0,\ldots,k_i}$$

$$= \rho^{\sum_{j=0}^{i} \ell_j k_j} \cdot \tau_{k_0,\ldots,k_i} .$$

The minimal polynomial of $\xi_{i+1}^{1/q}$ over K is $X^q - \xi_{i+1}$. Hence the conjugates of $\xi_{i+1}^{1/q}$ are $\rho^j \cdot \xi_{i+1}^{1/q}$ for $j = 0, 1, \ldots, q-1$, all with equal multiplicity. There exist numbers $m_j \in \{ 0, 1, \ldots, q-1 \}$ such that for $j = 0, 1, \ldots, q-1$ we have

$$\sigma_j(\xi_{i+1}^{1/q}) = \rho^{m_j} \cdot \xi_{i+1}^{1/q} .$$

Hence

$$\varphi_{\ell_0,\ldots,\ell_i}(\xi_{i+1}^{1/q}) = \rho^{\sum_{j=0}^{i} \ell_j m_j} \cdot \xi_{i+1}^{1/q} .$$

Now apply $\varphi_{\ell_0,\ldots,\ell_i}$ to (7.19). Then for each tuple $(\ell_0,\ldots,\ell_i)$ we find

$$\rho^{\sum_{j=0}^{i} \ell_j m_j} \cdot \xi_{i+1}^{1/q} = \sum_{k_0,\ldots,k_i} a_{k_0,\ldots,k_i} \cdot \rho^{\sum_{j=0}^{i} \ell_j k_j} \cdot \tau_{k_0,\ldots,k_i} .$$

Here we have a system of q^{i+1} linear equations in the q^{i+1} unknowns $a_{k_0,\ldots,k_i}$. The determinant of this system is exactly the square root of the discriminant of K_i over K , hence nonzero. Consequently there is in $\mathbb{C}^{q^{i+1}}$ just one solution of the system. But we know that solution:

$$a_{k_0,\ldots,k_i} = 0 \quad \text{if} \quad (k_0,\ldots,k_i) \neq (m_0,\ldots,m_i) ,$$
$$a_{m_0,\ldots,m_i} = \xi_{i+1}^{1/q} \cdot \tau_{m_0,\ldots,m_i}^{-1} .$$

The latter equation now yields an equation over K :

$$\xi_{i+1} = a_{m_0,\ldots,m_i}^q \cdot \prod_{j=0}^{i} \xi_j^{m_j} .$$

In case (1) this leads to the ideal equation

$$\left(\frac{\mathfrak{p}_{i+1}}{\mathfrak{p}'_{i+1}}\right)^{h_{i+1}} = \mathfrak{a}^q \cdot \prod_{j=1}^{i} \left(\frac{\mathfrak{p}_j}{\mathfrak{p}'_j}\right)^{m_j \cdot h_j} ,$$

and in case (2) to

$$\mathfrak{p}_{i+1}^{(\prime)\,h_{i+1}} = \mathfrak{a}^q \cdot \prod_{j=1}^{i} \mathfrak{p}_j^{(\prime)\,m_j \cdot h_j} ,$$

(where $\mathfrak{p}^{(\prime)}$ stands for $\mathfrak{p}$ or $\mathfrak{p}'$) for some fractional ideal $\mathfrak{a}$ (note that $(\xi_0) = (1)$). Because of unique factorization for ideals it follows in both cases that q divides all $m_j \cdot h_j$ for $j = 1, \ldots, i$ and h_{i+1} . This contradicts the assumption $q \nmid h$. □

Remarks. 1. If $\mathrm{ord}_p(\alpha_1^{b_1} \cdot \ldots \cdot \alpha_n^{b_n} - 1) > 1/(p-1)$ then

$$\mathrm{ord}_p(\alpha_1^{b_1} \cdot \ldots \cdot \alpha_n^{b_n} - 1) = \mathrm{ord}_p(b_1 \cdot \log_p(\alpha_1) + \ldots + b_n \cdot \log_p(\alpha_n)) .$$

We prefer to work with the logarithmic version, since that is the one we use in the computational method of reducing the upper bounds.

2. In order to apply Yu's lemma we can take for q the smallest odd prime that does not divide $h \cdot p \cdot (p^{f_{\mathfrak{p}}} - 1)$.

3. The author is grateful to M.A.J.G. van der Vlugt (Leiden) for discussions on the above lemma.

We now proceed to compute the constants C_1 to C_{12} . To find C_1 and C_2 we apply Lemma 2.6 to Λ_i^* , for all $i \in I_U$. Then we find for each such i constants $C_{1,i}$, $C_{2,i}$ such that, under the conditions

$$u_i + \lambda_i \geqslant \gamma_i , \quad B^* \geqslant \max \left(2, \frac{4}{3} \cdot t_i \cdot (p_i^{f_{\mathfrak{p}_i}/2} - 1) \right) ,$$

(where t_i denotes the number of terms in Λ_i^*), we obtain

$$\mathrm{ord}_{\mathfrak{p}_i}(\Lambda_i^*) < C_{1,i} + C_{2,i} \cdot \log B^* .$$

By Lemma 7.6(i) and the relation $\mathrm{ord}_{\mathfrak{p}} = e_{\mathfrak{p}} \cdot \mathrm{ord}_p$, and assuming that

$$U \geqslant \max_{i \in I_U} (\gamma_i - \lambda_i) , \quad B^* \geqslant \max_{i \in I_U} \left(2, \frac{4}{3} \cdot t_i \cdot (p_i^{f_{\mathfrak{p}_i}/2} - 1) \right) , \tag{7.20}$$

we see that it suffices to take

$$C_1 = \max_{i \in I_U} \left(-(\lambda_i + \mathrm{ord}_{p_i}(h^*)) + C_{1,i}/e_{\mathfrak{p}_i} \right) , \quad C_2 = \max_{i \in I_U} (C_{2,i}/e_{\mathfrak{p}_i}) .$$

Then (7.14) holds.

Next we apply Lemma 2.6 to K_i^* and $K_i'^*$, for all $i \in I$ and I' respectively, to obtain C_3 and C_4. By $X^{(')}$ we denote X if $i \in I$, and X' if $i \in I'$. There exist by Lemma 2.6 constants $C_{3,i}$ and $C_{4,i}$ such that under the conditions

$$h_i \cdot c_i + \kappa_i^{(')} \geqslant \gamma_i \ , \quad B^* \geqslant \max\left(2, \frac{4}{3} \cdot t_i \cdot (p_i^{f_{\mathfrak{p}_i}/2} - 1)\right)$$

(where again t_i denotes the number of terms of $K_i^{(')*}$), it follows that

$$\mathrm{ord}_{\mathfrak{p}_i}(K_i^{(')*}) < C_{3,i} + C_{4,i} \cdot \log B^* \ .$$

Again, by Lemma 7.6(ii),(ii') it follows that, under the conditions

$$M \geqslant \max_{i \in I \cup I'} \left(\frac{\gamma_i - \kappa_i^{(')}}{h_i}\right) \ , \quad B^* \geqslant \max_{i \in I \cup I'} \left(2, \frac{4}{3} \cdot t_i \cdot (p_i^{f_{\mathfrak{p}_i}/2} - 1)\right) \tag{7.21}$$

it suffices to take

$$C_3 = \max_{i \in I \cup I'} \left(\frac{\kappa_i^{(')} + \mathrm{ord}_{\mathfrak{p}_i}(h^*)}{h_i} + \frac{C_{3,i}}{h_i \cdot e_{\mathfrak{p}_i}} \right) \ , \quad C_4 = \max_{i \in I \cup I'} \left(\frac{C_{4,i}}{h_i \cdot e_{\mathfrak{p}_i}}\right) \ .$$

Then (7.15) holds.

We take C_5 to C_7 as follows:

$$C_5 = \log\left(2 \cdot \left|\frac{\alpha'}{\alpha}\right|\right) / 2 \cdot \log \varepsilon \ , \quad C_6 = \log\left(2 \cdot \left|\frac{\alpha}{\alpha'}\right|\right) / 2 \cdot \log \varepsilon \ ,$$

$$C_7 = \left(\sum_{i \in I} \log\left|\frac{\pi_i}{\pi_i'}\right| + \sum_{i \in I'} \log\left|\frac{\pi_i'}{\pi_i}\right| \right) / 2 \cdot \log \varepsilon \ .$$

Note that C_5 or C_6 may be negative, but that always $-C_6 < C_5$. Further, C_7 is always strictly positive, unless $I = I' = \emptyset$. Next we show how to take C_8 and C_9. Suppose first that

$$n > \max(C_5, 0) \ .$$

Then, from $\varepsilon \cdot \varepsilon' = \pm 1$ and the choice of π_i we find by (7.8) that

$$\left|\frac{\chi}{\chi'}\right| = \left|\frac{\alpha}{\alpha'}\right| \cdot \left|\frac{\varepsilon}{\varepsilon'}\right|^n \cdot \prod_{i \in I} \left|\frac{\pi_i}{\pi_i'}\right|^{c_i} \cdot \prod_{i \in I'} \left|\frac{\pi_i'}{\pi_i}\right|^{c_i} \geqslant \left|\frac{\alpha}{\alpha'}\right| \cdot \varepsilon^{2 \cdot n} > 2 \ ,$$

which expresses that the first term of G_α dominates. Put

$$P = \prod_{i\in I_U} p_i \ .$$

Then we infer

$$P^U \geqslant \prod_{i\in I_U} p_i^{\,u_i} = |\chi-\chi'|/2\cdot\sqrt{D} > |\chi|/4\cdot\sqrt{D}$$

$$= \frac{|\alpha|}{4\sqrt{D}}\cdot\varepsilon^n\cdot\prod_{i\in I}|\pi_i|^{c_i}\cdot\prod_{i\in I'}|\pi'_i|^{c_i} > \frac{|\alpha|}{4\sqrt{D}}\cdot\varepsilon^n \ ,$$

hence

$$n < \Big(\log\big(\frac{4\sqrt{D}}{|\alpha|}\big) + U\cdot\log(P)\Big)/\log\varepsilon \ .$$

Next suppose that

$$n < \min\ (\ -(C_6+C_7\cdot M),\ 0\ )\ .$$

Then we find that the second term of G_α dominates, namely

$$\left|\frac{\chi'}{\chi}\right| = \left|\frac{\alpha'}{\alpha}\right|\cdot\left|\frac{\varepsilon'}{\varepsilon}\right|^n\cdot\prod_{i\in I}\left|\frac{\pi'_i}{\pi_i}\right|^{c_i}\cdot\prod_{i\in I'}\left|\frac{\pi_i}{\pi'_i}\right|^{c_i}$$

$$\geqslant \left|\frac{\alpha'}{\alpha}\right|\cdot\varepsilon^{-2\cdot n}\cdot\left(\prod_{i\in I}\left|\frac{\pi'_i}{\pi_i}\right|\cdot\prod_{i\in I'}\left|\frac{\pi_i}{\pi'_i}\right|\right)^M = \left|\frac{\alpha'}{\alpha}\right|\cdot\varepsilon^{-2\cdot(n+C_7\cdot M)}$$

$$> \left|\frac{\alpha'}{\alpha}\right|\cdot\varepsilon^{2\cdot C_6} = 2\ .$$

Put

$$\Gamma = \prod_{i\in I}\min\ (\ 1,\ |\pi'_i|\ )\ \cdot\ \prod_{i\in I'}\min\ (\ 1,\ |\pi_i|\ )\ .$$

Then we infer

$$P^U \geqslant |\chi-\chi'|/2\cdot\sqrt{D} > |\chi'|/4\cdot\sqrt{D} = \frac{|\alpha'|}{4\sqrt{D}}\cdot\varepsilon^{|n|}\cdot\prod_{i\in I}|\pi'_i|^{c_i}\cdot\prod_{i\in I'}|\pi_i|^{c_i}$$

$$\geqslant \frac{|\alpha'|}{4\sqrt{D}}\cdot\varepsilon^{|n|}\cdot\prod_{i\in I}\min(1,|\pi'_i|)^{c_i}\cdot\prod_{i\in I'}\min(1,|\pi_i|)^{c_i}$$

$$\geqslant \frac{|\alpha'|}{4\sqrt{D}}\cdot\varepsilon^{|n|}\cdot\Gamma^M > \frac{|\alpha'|}{4\sqrt{D}}\cdot\varepsilon^{|n|}\cdot\Gamma^{-(|n|-C_6)/C_7}\ .$$

Hence

$$|n| < \Big(\log\big(\frac{4\sqrt{D}}{|\alpha'|}\cdot\Gamma^{-C_6/C_7}\big) + U\cdot\log(P) \Big)/\log\big(\varepsilon\cdot\Gamma^{1/C_7}\big) .$$

The remaining possibilities in cases (b) and (c) are $C_5 < n \leqslant 0$ and $0 \leqslant n < -(C_6+C_7\cdot M) < -C_6$. So we may take, noting that $\Gamma \leqslant 1$,

$$C_8 = \max \Big(\log\big(\frac{4\sqrt{D}}{|\alpha|}\big)/\log\,\varepsilon,\ \log\big(\frac{4\sqrt{D}}{|\alpha'|}\cdot\Gamma^{-C_6/C_7}\big)/\log\big(\varepsilon\cdot\Gamma^{1/C_7}\big),\ -C_5,\ -C_6 \Big) ,$$

$$C_9 = (\log P)/\log\big(\varepsilon\cdot\Gamma^{1/C_7}\big) .$$

Then (7.18) holds in the cases (b) and (c). Now take

$$C_{10} = \max \big(C_1,\ C_3,\ |C_5|,\ |C_6|+C_3\cdot C_7,\ C_8+C_1\cdot C_9 \big) ,$$

$$C_{11} = \max \big(C_2,\ C_4,\ C_4\cdot C_7,\ C_2\cdot C_9 \big) .$$

Then it follows that (7.16) is true, if conditions (7.20) and (7.21) hold. Hence, by Lemma 2.1, we infer the following result.

LEMMA 7.8. *In the above notation,*

$$B^* < C_{12}^* , \quad B < C_{12}$$

hold unconditionally, where

$$C_{12}^* = \max \Big(2\cdot\big(N+h^*\cdot C_{10}+h^*\cdot C_{11}\cdot\log(h^*\cdot C_{11})\big),\ \max_{i\in I_U} \big(h^*\cdot(\gamma_i-\lambda_i)+N\big),$$

$$\max_{i\in I\cup I'} \big(h^*\cdot\frac{\gamma_i-\kappa_i^{(\prime)}}{h_i}+N\big),\ 2,\ \max_{i\in I\cup I'\cup I_U} \big(\tfrac{4}{3}\cdot t_i\cdot(p_i^{f_{\mathfrak{p}_i}/2}-1)\big) \Big) ,$$

$$C_{12} = \frac{1}{h^*}\cdot(C_{12}^*+N) .$$

Proof. Clear. □

Remarks. 1. Theorem 7.1 is an immediate corollary of Lemma 7.8.
2. In practice, almost always the first term in the max-definition of C_{12}^* dominates. Moreover, the term N will in practice disappear in the rounding off. Similarly, in the definitions of C_{10} and C_{11} , the dominating factors are in practice C_1 to C_4 .

7.7. The reduction technique.

We now want to reduce the upper bound C_{12} for B (or C_{12}^{*} for B^{*}, which is equivalent), to a much smaller upper bound. We do so using the p-adic computational diophantine approximation technique described in Section 3.11.

We perform this procedure for $\Lambda = \Lambda_i^{*}$, K_i^{*}, $K_i'^{*}$, for the relevant i. We work in the p-adic approximation lattices Γ_μ themselves, and not in the sublattices described in Section 3.13. The computational bottlenecks are the computation of the p-adic logarithms to the desired precision, and the application of the L^3-Algorithm. We refer to Chapter 3 for details. Once we have found reduced bounds for $ord_p(\Lambda)$ for the above mentioned Λ, we combine these bounds with Lemma 7.6 and with estimates (7.13), (7.17) and (7.18) to find reduced bounds for B and B^{*}.

When reduced upper bounds for B, B^{*} are found in this way, we may try the above procedure again, with C_{12}, C_{12}^{*} replaced by their reduced analogons. We may repeat the argument as long as improvement is still being made. But at a certain stage, usually near to the actual largest solution, the procedure will not yield any further improvement. Then we have to find all solutions by some other method. One technique that may be useful is the algorithm of Fincke and Pohst, described in Section 3.6. Another way is to search directly for solutions of the original diophantine equation below the reduced bounds. In our present equation this may well be done by employing congruence arguments for finding all solutions of the second equation of system (7.9) below the obtained bounds.

7.8. The standard example.

In this section we shall work out the procedure outlined above for our standard example $\{ p_1, \ldots, p_s \} = \{ 2, 3, 5, 7 \}$, thus proving Theorem 7.2. In Tables II and III we give the necessary data on the fields $K = \mathbb{Q}(\sqrt{D})$ for the 15 values of D, and on the factorization of 2, 3, 5, 7 in K.

Explanation of Tables II and III. For $p_i = 2, 3, 5, 7$ we give in Table II a generator of the ideal $\mathfrak{p}_i$ with $ord_{p_i}(\mathfrak{p}_i) > 0$ if $\mathfrak{p}_i$ is a principal ideal, and we give "$\mathfrak{p}_i$" if it is not principal. In all the latter cases, $h_i = 2$, so $\mathfrak{p}_i^2 = (\pi_i)$ is principal. An asterisk (*) denotes a splitting prime. Note that for each D at most one of the primes $2, 3, 5, 7$ splits, so $t \leqslant 1$. In the final column of Table II we give for the splitting prime p_i a generator π_i of the ideal $\mathfrak{p}_i^{h_i}$. In Table III, when $\mathfrak{p}_i$ and $\mathfrak{p}_j$ are not principal, but $\mathfrak{p}_i \cdot \mathfrak{p}_j$ is, we give a generator of it. The autor is grateful to R.J. Kooman (Leiden) for checking these tables.

From Tables II and III it is easy to find all possibilities for I, I' and α. We may assume $I' = \emptyset$. In Table IV we give all possible I, I_U, α (we give primes p_i instead of indices i). An asterisk (*) appears when $(\alpha) \neq (\alpha')$. The set I_U is found by checking $G_\alpha \pmod{p_i}$ for all p_i.

There are 54 cases with $I = \emptyset$ (the "symmetric" cases), and 54 cases with $I \neq \emptyset$ (the "asymmetric" cases). We start with the symmetric cases. This incorporates all cases with $D = 3, 5, 35, 42, 210$, when none of the primes $2, 3, 5, 7$ splits in $\mathbb{Q}(\sqrt{D})$. Now, $t = 0$, hence equation (7.10) becomes

$$G_\alpha(n) = \frac{\alpha}{2\sqrt{D}} \cdot \varepsilon^n - \frac{\alpha'}{2\sqrt{D}} \cdot \varepsilon'^n = \pm \prod_{i \in I_U} p_i^{u_i} . \tag{7.22}$$

With $A = \varepsilon + \varepsilon' \in \mathbb{Z}$, $B = N\varepsilon = \varepsilon \cdot \varepsilon' = \pm 1$, we have for all $n \in \mathbb{Z}$

$$G_\alpha(n+2) = A \cdot G_\alpha(n+1) - B \cdot G_\alpha(n) .$$

Since $(\alpha) = (\alpha')$, there is an $n_0 \in \mathbb{Z}$ such that $\alpha' = \pm\varepsilon^{n_0} \cdot \alpha$. Hence

$$|G_\alpha(n_0 - n)| = |G_\alpha(n)|$$

for all $n \in \mathbb{Z}$, which explains why we call these cases "symmetric". In this situation we can apply elementary congruence arguments, as explained in Section 4.5. We have the following result.

LEMMA 7.9. *Let* $\{ p_1, \ldots, p_4 \} = \{ 2, 3, 5, 7 \}$. *Equation* (7.1) *with conditions* (7.2) *and* $I = \emptyset$ *has exactly* 91 *solutions, that appear in* Table I *marked with an asterisk* (*).

Sketch of proof. In Table V we give the necessary data for these 54 cases. We explain this table, and leave many details to the reader to check. For each $p = 2, 3, 5, 7$ we give $\ell_1, n_1, a_1, h_2, \ldots, h_7$. If for a p only ℓ_1 is given, then $p^{\ell_1+1} \nmid G_\alpha(n)$ for all $n \in \mathbb{Z}$, and $p^{\ell_1} \mid G_\alpha(n)$ for at least one $n \in \mathbb{Z}$. If n_1, a_1 are given, then

$$p^{\ell_1+1} \mid G_\alpha(n) \Leftrightarrow n \equiv n_1 \pmod{a_1} .$$

Define $n_2 = a_1$ if $n_1 = 0$, and $n_2 = n_1$ if $n_1 \neq 0$. Then n_2 is the smallest positive index such that $p^{\ell_1+1} \mid G_\alpha(n_2)$. Now it is true that

$$G_\alpha(n_2) \mid G_\alpha(n) \quad \text{whenever} \quad n \equiv n_1 \pmod{a_1} ,$$

This is related to symmetry properties of the recurrence sequence $\{G_\alpha(n)\}_{n=-\infty}^{\infty}$. For $q = 2, 3, 5, 7$ we have defined

$$h_q = \mathrm{ord}_q(G_\alpha(n_2)) .$$

Hence $2^{h_2}\cdot 3^{h_3}\cdot 5^{h_5}\cdot 7^{h_7} \mid G_\alpha(n)$ whenever $p^{\ell_1+1} \mid G_\alpha(n)$. We have taken ℓ_1 so large that always

$$G_\alpha(n_2) > 2^{h_2}\cdot 3^{h_3}\cdot 5^{h_5}\cdot 7^{h_7} . \tag{7.23}$$

Consequently, there exists some prime $r \geqslant 11$ that divides $G_\alpha(n_2)$, hence r divides all $G_\alpha(n)$ with $p^{\ell_1+1} \mid G_\alpha(n)$. It follows that for a solution of equation (7.22) we must have

$$\mathrm{ord}_p(G_\alpha(n)) \leqslant \ell_1 .$$

In this way we find with ease all solutions of (7.22). □

Let us illustrate this with the example $D = 3,\ \alpha = \sqrt{3}$. Then

$$G_\alpha(n) = \tfrac{1}{2}\cdot(2+\sqrt{3})^n + \tfrac{1}{2}\cdot(2-\sqrt{3})^n ,$$

and $G_\alpha(-n) = G_\alpha(n)$. We have for $G_\alpha(n)$:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$G_\alpha(n)$	1	2	7	26	97	362	....					$G_\alpha(14) = 50843527$				
mod 4	1	2	−1	2	1	2	−1	2	1	2	−1	2	1	2	−1	2
mod 3	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	−1
mod 5	1	2	2	1	2	2	1	2	2	1	2	2	1	2	2	1
mod 49	1	2	7	−23	−1	19	−21	−5	1	9	−14	−16	−1	12	0	−12

We see that $2^2, 3, 5 \nmid G_\alpha(n)$ for all $n \in \mathbb{Z}$, and $2 \mid G_\alpha(n)$ if and only if n odd . So $p = 7$ is the only interesting case. We have $7 \mid G_\alpha(n)$ if and only if $n \equiv 2 \pmod 4$, $7^2 \mid G_\alpha(n)$ if and only if $n \equiv 14 \pmod{28}$, (and in general

$$7^k \mid G_\alpha(n) \Leftrightarrow n \equiv 2\cdot 7^{k-1} \pmod{4\cdot 7^{k-1}})$$

for $k \geqslant 1$, and a similar relation holds for any symmetric recurrence and any prime p for which arbitrary high powers of p occur in $G_\alpha(n)$, cf. Lemma 4.10). Now, $\ell_1 = 0$ does not lead to (7.23), since then $n_2 = 2$, and $G_\alpha(2) = 7$, so that no suitable r exists. But with $\ell_1 = 1$ we have $n_2 = 14$, and $h_2 = h_3 = h_5 = 0$, $h_7 = 2$, and (7.23) holds, since $G_\alpha(14) > 7^2$. Hence there exists a prime $r \geqslant 11$ such that $r \mid G_\alpha(14)$, and thus $r \mid G_\alpha(n)$ whenever $7^2 \mid G_\alpha(n)$. It follows that for solutions of (7.22) we have $G_\alpha(n) \leqslant 2^1\cdot 3^0\cdot 5^0\cdot 7^1 = 14$, so that all solutions can be read from the above table. Note that it is not necessary that r is known explicitly, only that $G_\alpha(n_2)$ is large enough. In our example, $r = 337$ or $r = 3079$ satisfy.

Finally we treat the remaining 54 cases, where $I \neq \emptyset$. Then we need the non-elementary reduction technique described in Sections 7.5 to 7.7.

In all our instances, the set I contains only one element, since there is only one splitting prime. We denote by π the π_i belonging to this prime, and we write m for c_i . Equation (7.10) now reads

$$\frac{\alpha}{2\sqrt{D}}\cdot\varepsilon^n\cdot\pi^m - \frac{\alpha'}{2\sqrt{D}}\cdot\varepsilon'^n\cdot\pi'^m = \pm\prod_{j\in I_U} p_j^{u_j} .$$

We computed the constants C_1 to C_{12} , C_{12}^* , according to Section 7.6, for each of the 54 cases. We omit the details of these computations, and simply give the data in Table VI. In this table we give for each D the $p_i \in I_U$ together with the ν_i and λ_i (it turns out that the λ_i do not depend on the α , only on the p_i). The values "n_ε, n_π, n_2, n_3, n_5, n_7" are the integers such that

$$\alpha^2 = \pm\, \varepsilon^{n_\varepsilon}\cdot\pi^{n_\pi}\cdot 2^{n_2}\cdot\ldots\cdot 7^{n_7} .$$

It follows that in all cases we have $C_{12}^* < 3.23\times 10^{30}$.

The next step is to define the lattices, and find lower bounds for the shortest nonzero vectors in the lattices. We start with treating the Λ_i^* , of which there are 3 for each of the 10 D's . We have computed the 30 values of

$$\vartheta = -\frac{\log_{p_i}\left(\frac{\pi}{\pi'}\right)}{\log_{p_i}\left(\frac{\varepsilon}{\varepsilon'}\right)} \quad \text{or} \quad -\frac{\log_{p_i}\left(\frac{\varepsilon}{\varepsilon'}\right)}{\log_{p_i}\left(\frac{\pi}{\pi'}\right)} ,$$

such that it is a p_i-adic integer, to the desired precision of μ digits. We took μ as follows:

p_i	μ	p_i^{μ}
2	209	8.22×10^{62}
3	133	2.87×10^{63}
5	95	2.52×10^{66}
7	76	1.69×10^{64}

in order to have p_i^{μ} somewhat larger than the maximal C_{12}^{*2} , being 1.05×10^{61} . We computed the 30 values of the $\vartheta^{(\mu)}$'s , but do not give them here. The lattices Γ_{μ} are generated by the column vectors of the matrices

$$\begin{pmatrix} 1 & 0 \\ \vartheta^{(\mu)} & p^{\mu} \end{pmatrix} .$$

We performed the p-adic continued fraction algorithm of Section 3.10 for each of these 30 lattices. In the table below we give for each D the maximal C_{12}^{*} (there is one for each α), and the minimal bound for $\ell(\Gamma_{\mu})$ (there is one for each $i \in I_U$) that we found. We omit further details.

D	p	μ_0	$C_{12}^{*} \leqslant$	$\ell(\Gamma_{\mu}) >$	$U \leqslant$
2	2, 3, 5	1.5, 1.0, 1.0	3.19×10^{28}	8.26×10^{30}	210
6	2, 3, 7	1.5, 1.5, 1.0	2.72×10^{26}	2.05×10^{31}	210
7	2, 5, 7	2.0, 1.0, 0.5	1.07×10^{30}	2.43×10^{31}	210
10	2, 5, 7	1.5, 0.5, 1.0	3.22×10^{29}	2.22×10^{31}	210
14	2, 3, 7	1.5, 1.0, 0.5	4.80×10^{26}	1.48×10^{31}	210
15	2, 3, 5	3.5, 1.5, 0.5	2.15×10^{28}	1.55×10^{31}	212
21	2, 3, 7	3.0, 0.5, 0.5	1.90×10^{26}	7.78×10^{30}	211
30	2, 3, 5	2.5, 0.5, 0.5	4.15×10^{28}	1.37×10^{31}	211
70	2, 5, 7	2.5, 0.5, 0.5	3.23×10^{30}	2.51×10^{31}	211
105	3, 5, 7	1.5, 0.5, 0.5	4.54×10^{29}	3.96×10^{31}	134

In all cases, $\ell(\Gamma_{\mu}) > \sqrt{2}\cdot C_{12}^{*}$. Hence Lemma 3.14 with $n = 2$, $c_1 = 0$, $c_2 = 1$ yields

$$\mathrm{ord}_{p_i}(\Lambda_i^*) < \mu + \mu_0 \ , \quad i \in I_U \ ,$$

where

$$\mu_0 = \min \left(\mathrm{ord}_{p_i}(\log_{p_i}(\frac{\varepsilon}{\varepsilon'})), \ \mathrm{ord}_{p_i}(\log_{p_i}(\frac{\pi}{\pi'})) \right) \ ,$$

as given above. By $\lambda_i + \mathrm{ord}_{p_i}(h^*) \geqslant 0$ we obtain from Lemma 7.6(i) upper bounds for u_i , $i \in I_U$, hence the upper bounds for U , as given above.

Next, we treat the K_i^* , one for each D , having 5 terms, namely

$$K_i^* = n^* \cdot \log_{p_i}(\varepsilon') + m^* \cdot \log_{p_i}(\pi') - \sum_{\substack{1 \leqslant j \leqslant 4 \\ j \neq i}} u_j^* \cdot \log_{p_i}(p_j) \ ,$$

where $i \in I$, so p_i is the splitting prime. We have the following data.

D	p_i	$\sqrt{D} \pmod{p_i}$	$\mathrm{ord}_{p_i}(\log_{p_i}(\cdot))$					
			ε'	π'	2	3	5	7
2	7	3	1	2	1	1	1	-
6	5	4	1	1	1	1	-	2
7	3	1	1	1	1	-	1	1
10	3	2	1	1	1	-	1	1
14	5	2	1	1	1	1	-	2
15	7	6	1	1	1	1	1	-
21	5	4	1	1	1	1	-	2
30	7	4	1	1	1	1	1	-
70	3	2	1	1	1	-	1	1
105	2	1 (mod 4)	2	4	-	2	2	3

From this table our choice for $\sqrt{D} \pmod{p_i}$ becomes clear. It follows that $\mathrm{ord}_{p_i}(\log_{p_i}(\varepsilon'))$ is always the least one of the five ord_{p_i}'s in the above table. So we define:

$$\vartheta_1 = -\frac{\log_{p_i}(\pi')}{\log_{p_i}(\varepsilon')} \ , \quad \vartheta_{2,3,4} = -\frac{\log_{p_i}(p_j)}{\log_{p_i}(\varepsilon')} \ , \quad (j \in \{1,2,3,4\}, \ j \neq i) \ ,$$

and we computed these numbers up to μ digits, with μ as follows:

p_i	μ	p_i^{μ}
2	539	1.80×10^{162}
3	343	4.49×10^{163}
5	245	1.77×10^{171}
7	196	4.36×10^{165}

so that p_i^{μ} is somewhat larger than the maximal C_{12}^{*5}. We computed the 40 values of the $\vartheta_{1,2,3,4}^{(\mu)}$, but do not give them here. The lattices Γ_{μ} are generated by the columns of the following matrices:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \vartheta_1^{(\mu)} & \vartheta_2^{(\mu)} & \vartheta_3^{(\mu)} & \vartheta_4^{(\mu)} & p^{\mu} \end{pmatrix}.$$

We computed the reduced bases of the 10 lattices by the L^3-algorithm. Again, we omit the computational details. We found data as follows.

D	p in I	μ	μ_0	$C_{12}^{*} \leqslant$	$\ell(\Gamma_{\mu}) >$	$M \leqslant$
2	7	196	1	3.19×10^{28}	2.25×10^{32}	196
6	5	245	1	2.72×10^{26}	2.16×10^{33}	245
7	3	343	1	1.07×10^{30}	1.14×10^{32}	343
10	3	343	1	3.22×10^{29}	1.07×10^{32}	343
14	5	245	1	4.80×10^{26}	4.92×10^{33}	245
15	7	196	1	2.15×10^{28}	2.78×10^{32}	196
21	5	245	1	1.90×10^{26}	4.37×10^{33}	245
30	7	196	1	4.15×10^{28}	2.69×10^{32}	196
70	3	343	1	3.23×10^{30}	1.03×10^{32}	343
105	2	539	2	4.54×10^{29}	6.68×10^{31}	540

In all instances, $\ell(\Gamma_{\mu}) > \sqrt{5}\cdot C_{12}^{*}$, so that by Lemmas 3.14 and 7.6(ii) and $\kappa_i + \mathrm{ord}_{p_i}(h^{*}) \geqslant 0$ and $h_i \geqslant 1$ we have $M \leqslant \mathrm{ord}_{p_i}(K_i^{*}) < \mu + \mu_0$, hence an upper bound for M as given in the table above.

Finally, we compute the new, reduced bounds for $|n|$, and thus for B, by

$$|n| < \max\left(C_5,\ C_6 + C_7\cdot M,\ C_8 + C_9\cdot U \right).$$

Hence we find data as in the following table.

D	C_5 <	C_6 <	C_7 <	C_8 <	C_9 <	M ⩽	U ⩽	|n| ⩽	B ⩽	N ⩽	B^* ⩽
2	0.394	0.394	0.420	1.967	3.859	196	210	812	812	3	1627
6	0.152	0.652	0.190	1.345	1.631	245	210	343	343	3	689
7	0.126	0.626	0.357	2.702	2.757	343	210	581	581	2	1164
10	0.601	0.191	0.181	1.396	2.337	343	210	492	492	3	987
14	0.102	0.602	0.325	1.861	1.508	245	210	318	318	3	639
15	0.540	0.668	0.257	1.394	1.649	196	212	350	350	2	702
21	0.222	0.722	0.142	1.564	2.386	245	211	505	505	1	1011
30	0.414	0.613	0.399	1.239	1.102	196	211	233	233	3	469
70	0.362	0.556	0.390	2.729	1.505	343	211	320	343	3	689
105	0.390	0.579	0.379	3.232	2.545	540	134	344	540	1	1081

Here we used $B^* \leqslant h^* \cdot B + N$ and $h^* = 2$. So in one step we have reduced the bound $B^* < 3.23 \times 10^{30}$ to $B^* \leqslant 1627$. The total computation time was 1715 sec, on average 0.7 sec for each 2-dimensional lattice, and 170 sec for each 5-dimensional lattice.

We made a further reduction step, now using the reduced bound for B^* as given above in stead of C_{12}^* . We give the data for the Λ_i^* in the tables below. For μ we took $\mu_1 \cdot \mu_2$, with μ_1, μ_2 as below:

p	2	3	5	7
μ_2	11	7	5	4

,

D	B^* ⩽	$\sqrt{2} \cdot B^*$ <	μ_1	μ ⩽	$\ell(\Gamma_\mu)$ ⩾	μ_0 ⩽	U ⩽
2	1627	2301	2	22	1.82×10^3	1.5	23
6	689	975	3	33	3.99×10^4	1.5	34
7	1164	1647	3	33	4.50×10^4	2	34
10	987	1396	3	33	5.91×10^4	1.5	34
14	639	904	3	33	2.58×10^4	1.5	34
15	702	993	3	33	7.36×10^4	3.5	36
21	1011	1430	3	33	2.00×10^4	3	35
30	469	664	2	22	9.98×10^2	2.5	24
70	689	975	3	33	5.76×10^4	2.5	35
105	1081	1529	3	21	3.89×10^4	1.5	22

We found $\ell(\Gamma_\mu)$ and bounds for U as given in the above table. For the K_i^* we found, with $\mu = \mu_1 \cdot \mu_2$ with μ_2 as above, and μ_1 as in the table below, the results given in that table.

D	$B^* \leqslant$	$\sqrt{5} \cdot B^* <$	μ_1	$\mu \leqslant$	$\ell(\Gamma_\mu) \geqslant$	$\mu_0 \leqslant$	$M \leqslant$	$\lvert n \rvert \leqslant$	$B \leqslant$	$B^* \leqslant$
2	1627	3639	7	28	1.24×10^4	1	28	90	90	183
6	689	1541	6	30	4.04×10^3	1	30	145	145	293
7	1164	2603	7	49	1.07×10^4	1	49	96	96	194
10	987	2207	7	49	1.16×10^4	1	49	80	80	163
14	639	1429	6	30	3.07×10^3	1	30	53	53	109
15	702	1570	6	24	2.70×10^3	1	24	60	60	122
21	1011	2261	6	30	3.88×10^3	1	30	85	85	171
30	469	1049	6	24	2.50×10^3	1	24	27	27	57
70	689	1541	6	42	1.90×10^3	1	42	55	55	113
105	1081	2418	7	77	1.00×10^4	2	78	59	78	157

The computation time was 15 sec.

We made a third step, and give data like above, for Λ_i^* :

D	$B^* \leqslant$	$\sqrt{2} \cdot B^* <$	μ_1	$\mu \leqslant$	$\ell(\Gamma_\mu) \geqslant$	$\mu_0 \leqslant$	$U \leqslant$
2	183	258.9	2	22	1821	1.5	23
6	299	414.4	2	22	875	1.5	23
7	194	274.4	2	22	1285	2	23
10	163	230.6	2	22	634	1.5	23
14	109	154.2	2	22	268	1.5	23
15	122	172.6	2	22	873	3.5	25
21	171	241.9	2	22	818	3	25
30	57	80.7	2	22	998	2.5	24
70	113	159.9	2	22	585	2.5	24
105	157	222.1	2	14	281	1.5	15

and for K_i^* :

D	$B^* \leqslant$	$\sqrt{5}\cdot B^* <$	μ_1	$\mu \leqslant$	$\ell(\Gamma_\mu) \geqslant$	$\mu_0 \leqslant$	$M \leqslant$
2	183	409.3	5	20	440	1	20
6	293	655.2	5	25	665	1	25
7	194	433.8	6	42	602	1	42
10	163	364.5	5	35	473	1	35
14	109	243.8	5	25	626	1	25
15	122	272.9	6	24	2700	1	24
21	171	382.4	5	25	645	1	25
30	57	127.5	4	16	129	1	16
70	113	252.7	5	35	366	1	35
105	157	351.1	5	55	354	2	56

and finally for $|n|$, and in more detail for $\mathrm{ord}_{p_i}(u)$ for $i \in I_U$

D	$M \leqslant$	$u_2 \leqslant$	$u_3 \leqslant$	$u_5 \leqslant$	$u_7 \leqslant$	$\|n\| \leqslant$
2	20	23	14	10	0	90
6	25	23	15	0	8	38
7	42	23	0	10	8	66
10	35	23	0	10	8	55
14	25	23	14	0	8	36
15	24	25	15	10	0	42
21	25	24	14	0	8	61
30	16	24	14	10	0	27
70	35	24	0	10	8	65
105	56	0	14	10	8	41

Now we will not find any further improvement if we proceed in the same way. But the upper bounds are now small enough to admit enumeration of the remaining possibilities, making use of mod p arithmetic for $p = 2, 3, 5, 7$. We did so, and found the remaining solutions, presented in Table I. We used only 3 sec computer time for this last step.

This completes the proof of Theorem 7.2. □

7.9. Tables.

Table I. (Theorem 7.2.)

Nr	X	Y	Z	D
1	4375	-4374	1	7
2	2401	-2400	1	1
3	225	-224	1	1
4	126	-125	1	14
5	81	-80	1	1
6	64	-63	1	1
7	50	-49	1	2
8	49	-48	1	1
9	36	-35	1	1
10	28	-27	1	7
11	25	-24	1	1
12	21	-20	1	21
13	16	-15	1	1
14	15	-14	1	15
15	10	-9	1	10
16	9	-8	1	1
17	8	-7	1	2
18	7	-6	1	7
19	6	-5	1	6
20	5	-4	1	5
21	4	-3	1	1
22	3	-2	1	3
23	2	-1	1	2
24	490	-486	2	10
25	54	-50	2	6
26	49	-45	2	1
27	25	-21	2	1
28	18	-14	2	2
29	14	-10	2	14
30	10	-6	3	10
31	9	-5	2	1
32	7	-3	2	7
33	6	-2	2	6
34	5	-1	2	5
35	3	1	2	3
36	2	2	2	2
37	384	-375	3	6
38	105	-96	3	105
39	84	-75	3	21
40	49	-40	3	1
41	30	-21	3	30
42	25	-16	3	1
43	24	-15	3	6
44	21	-12	3	21
45	16	-7	3	1
46	15	-6	3	15
47	14	-5	3	14
48	12	-3	3	3
49	10	-1	3	10
50	8	1	3	2
51	7	2	3	7
52	6	3	3	6
53	5	4	3	5
54	70	-54	4	70
55	30	-14	4	30
56	25	-9	4	1
57	21	-5	4	21
58	18	-2	4	2
59	15	1	4	15
60	14	2	4	14
61	10	6	4	10
62	9	7	4	1
63	5145	-5120	5	105
64	270	-245	5	30
65	160	-135	5	10
66	105	-80	5	105
67	81	-56	5	1
68	70	-45	5	70
69	60	-35	5	15
70	49	-24	5	1
71	45	-20	5	5
72	40	-15	5	10
73	35	-10	5	35
74	32	-7	5	2
75	30	-5	5	30
76	28	-3	5	7
77	27	-2	5	3
78	24	1	5	6
79	21	4	5	21
80	20	5	5	5
81	18	7	5	2
82	16	9	5	1
83	15	10	5	15
84	50	-14	6	2
85	42	-6	6	42
86	35	1	6	35
87	30	6	6	30
88	21	15	6	21
89	1750	-1701	7	70
90	945	-896	7	105
91	625	-576	7	1
92	224	-175	7	14
93	189	-140	7	21
94	175	-126	7	7
95	112	-63	7	7
96	105	-56	7	105
97	84	-35	7	21
98	81	-32	7	1
99	70	-21	7	70
100	64	-15	7	1

Table I. (cont.)

Nr	X	Y	Z	D	Nr	X	Y	Z	D
101	63	-14	7	7	151	72	49	11	2
102	56	-7	7	14	152	294	-150	12	6
103	54	-5	7	6	153	150	-6	12	6
104	50	-1	7	2	154	147	-3	12	3
105	48	1	7	3	155	729	-560	13	1
106	45	4	7	5	156	512	-343	13	2
107	42	7	7	42	157	294	-125	13	6
108	40	9	7	10	158	250	-81	13	10
109	35	14	7	35	159	225	-56	13	1
110	28	21	7	7	160	196	-27	13	1
111	25	24	7	1	161	189	-20	13	21
112	750	-686	8	30	162	175	-6	13	7
113	189	-125	8	21	163	168	1	13	42
114	162	-98	8	2	164	162	7	13	2
115	70	-6	8	70	165	160	9	13	10
116	63	1	8	7	166	144	25	13	1
117	54	10	8	6	167	120	49	13	30
118	50	14	8	2	168	105	64	13	105
119	49	15	8	1	169	250	-54	14	10
120	375	-294	9	15	170	210	-14	14	210
121	256	-175	9	1	171	189	7	14	21
122	105	-24	9	105	172	175	21	14	7
123	96	-15	9	6	173	126	70	14	14
124	84	-3	9	21	174	960	-735	15	15
125	80	1	9	5	175	245	-20	15	5
126	75	6	9	3	176	240	-15	15	15
127	60	21	9	15	177	224	1	15	14
128	56	25	9	14	178	210	15	15	210
129	49	32	9	1	179	120	105	15	30
130	343	-243	10	7	180	270	-14	16	30
131	135	-35	10	15	181	250	6	16	10
132	105	-5	10	105	182	175	81	16	7
133	98	2	10	2	183	6561	-6272	17	1
134	90	10	10	10	184	1024	-735	17	1
135	70	30	10	70	185	625	-336	17	1
136	625	-504	11	1	186	343	-54	17	7
137	441	-320	11	1	187	324	-35	17	1
138	256	-135	11	1	188	294	-5	17	6
139	196	-75	11	1	189	288	1	17	2
140	175	-54	11	7	190	280	9	17	70
141	135	-14	11	15	191	240	49	17	15
142	128	-7	11	2	192	225	64	17	1
143	126	-5	11	14	193	189	100	17	21
144	125	-4	11	5	194	294	30	18	6
145	120	1	11	30	195	1225	-864	19	1
146	112	9	11	7	196	486	-125	19	6
147	105	16	11	105	197	441	-80	19	1
148	100	21	11	1	198	375	-14	19	15
149	96	25	11	6	199	360	1	19	10
150	81	40	11	1	200	343	18	19	7

Table I. (cont.)

Nr	X	Y	Z	D
201	336	25	19	21
202	280	81	19	70
203	256	105	19	1
204	490	-90	20	10
205	405	-5	20	5
206	525	-84	21	21
207	448	-7	21	7
208	420	21	21	105
209	336	105	21	21
210	729	-245	22	1
211	490	-6	22	10
212	486	-2	22	6
213	1215	-686	23	15
214	1029	-500	23	21
215	729	-200	23	1
216	625	-96	23	1
217	525	4	23	21
218	504	25	23	14
219	480	49	23	30
220	448	81	23	7
221	625	-49	24	1
222	945	-320	25	105
223	640	-15	25	10
224	630	-5	25	70
225	576	49	25	1
226	490	135	25	10
227	686	-10	26	14
228	675	1	26	3
229	1029	-300	27	21
230	750	-21	27	30
231	735	-6	27	15
232	1134	-350	28	14
233	1225	-384	29	1
234	840	1	29	210
235	729	112	29	1
236	625	216	29	1
237	441	400	29	1
238	6561	-5600	31	1
239	2401	-1440	31	1
240	1024	-63	31	1
241	960	1	31	15
242	945	16	31	105
243	625	336	31	1
244	1029	-5	32	21
245	2625	-1536	33	105
246	1029	60	33	21
247	1792	-567	35	7
248	1260	-35	35	35
249	1215	10	35	15
250	1120	105	35	70
251	945	280	35	105
252	1372	-3	37	7
253	1344	25	37	21
254	1225	144	37	1
255	729	640	37	1
256	1458	-14	38	2
257	1536	-15	39	6
258	1500	21	39	15
259	896	625	39	14
260	2401	-720	41	1
261	1701	-20	41	21
262	1680	1	41	105
263	1600	81	41	1
264	1750	14	42	70
265	1800	49	43	2
266	1120	729	43	70
267	1250	686	44	2
268	1920	105	45	30
269	16384	-14175	47	1
270	2401	-192	47	1
271	2205	4	47	5
272	2160	49	47	15
273	2625	-224	49	105
274	2400	1	49	6
275	1701	700	49	21
276	2430	70	50	30
277	2625	-24	51	105
278	2401	200	51	1
279	15309	-12500	53	21
280	2800	9	53	7
281	2025	784	53	1
282	3430	-405	55	70
283	3024	1	55	21
284	3150	-14	56	14
285	3200	49	57	2
286	4050	-686	58	2
287	3456	25	59	6
288	2401	1080	59	1
289	35721	-32000	61	1
290	4096	-375	61	1
291	3969	-125	62	1
292	2625	1344	63	105
293	3969	256	65	1
294	4480	9	67	70
295	4374	250	68	6
296	5145	-384	69	105
297	15625	-10584	71	1
298	5040	1	71	35
299	4096	945	71	1
300	4704	625	73	6

Table I. (cont.)

Nr	X	Y	Z	D
301	5145	480	75	105
302	3375	2401	76	15
303	6804	-875	77	21
304	6561	-320	79	1
305	6250	-9	79	10
306	3840	2401	79	15
307	8505	-1280	85	105
308	7840	81	89	10
309	65625	-57344	91	105
310	8505	-224	91	105
311	10240	-1215	95	10
312	9408	1	97	3
313	9800	1	99	2
314	10206	-5	101	14
315	9375	1029	102	15
316	10584	25	103	6
317	11250	-14	106	2
318	12544	225	113	1
319	10368	2401	113	2
320	13230	-5	115	30
321	15625	-1701	118	1
322	14336	-175	119	14
323	14175	-14	119	7
324	14406	-6	120	6
325	18225	-3584	121	1
326	16128	1	127	7
327	15625	504	127	1
328	15625	1536	131	1
329	17500	189	133	7
330	18144	625	137	14
331	18750	294	138	30
332	117649	-97200	143	1
333	21504	105	147	21
334	24010	15	155	10
335	23625	1024	157	105
336	25920	1	161	5
337	26250	-6	162	42
338	16807	13122	173	7
339	30618	7	175	42
340	32768	-7	181	2
341	33614	-125	183	14
342	43740	-1715	205	15
343	43750	-486	208	70
344	46305	-80	215	105
345	50625	-896	223	1
346	49000	729	223	10
347	129654	-78125	227	6
348	55566	-3125	229	14
349	60025	-1944	241	1
350	59535	1	244	15
351	59049	1960	247	1
352	63000	1	251	70
353	64000	9	253	10
354	48384	15625	253	21
355	59049	7000	257	1
356	69120	49	263	30
357	85750	-486	292	70
358	83349	2500	293	21
359	140625	-43904	311	1
360	109375	-1134	329	7
361	82944	30625	337	1
362	128625	256	359	105
363	137781	-140	371	21
364	76545	71680	385	105
365	196830	-33614	404	30
366	117649	48000	407	1
367	168070	30	410	70
368	179200	6561	431	7
369	137200	59049	443	7
370	201684	-1875	447	21
371	201600	1	449	14
372	214375	-6	463	7
373	252105	-24576	477	105
374	243000	49	493	30
375	245760	-735	495	15
376	262144	5145	517	1
377	390625	-112896	527	1
378	688905	-5	830	105
379	1058841	-20480	1019	1
380	1440000	2401	1201	1
381	1640625	336	1281	105
382	4214784	25	2053	21
383	4782969	4375	2188	1
384	5764801	-9600	2399	1
385	19140625	-17496	4373	1
386	23049600	1	4801	6
387	76545000	1	8749	42
388	199290375	-686	14117	15

Table II.

D	h	ε	$N\varepsilon$	$\mathfrak{p}_1$	$\mathfrak{p}_2$	$\mathfrak{p}_3$	$\mathfrak{p}_4$	π_i
2	1	$1+\sqrt{2}$	-1	$\sqrt{2}$	3	5	$1+2\sqrt{2}^*$	$1+2\sqrt{2}$
3	1	$2+\sqrt{3}$	1	$1+\sqrt{3}$	$\sqrt{3}$	5	7	–
5	1	$\frac{1}{2}(1+\sqrt{5})$	-1	2	3	$\sqrt{5}$	7	–
6	1	$5+2\sqrt{6}$	1	$2+\sqrt{6}$	$3+\sqrt{6}$	$1+\sqrt{6}^*$	7	$1+\sqrt{6}$
7	1	$8+3\sqrt{7}$	1	$3+\sqrt{7}$	$2+\sqrt{7}^*$	5	$\sqrt{7}$	$2+\sqrt{7}$
10	2	$3+\sqrt{10}$	-1	$\mathfrak{p}_1$	$\mathfrak{p}_2^*$	$\mathfrak{p}_3$	7	$1+\sqrt{10}$
14	1	$15+4\sqrt{14}$	1	$4+\sqrt{14}$	3	$3+\sqrt{14}^*$	$7+2\sqrt{14}$	$3+\sqrt{14}$
15	2	$4+\sqrt{15}$	1	$\mathfrak{p}_1$	$\mathfrak{p}_2$	$\mathfrak{p}_3$	$\mathfrak{p}_4^*$	$8+\sqrt{15}$
21	1	$\frac{1}{2}(5+\sqrt{21})$	1	2	$\frac{1}{2}(3+\sqrt{21})$	$\frac{1}{2}(1+\sqrt{21})^*$	$\frac{1}{2}(7+\sqrt{21})$	$\frac{1}{2}(1+\sqrt{21})$
30	2	$11+2\sqrt{30}$	1	$\mathfrak{p}_1$	$\mathfrak{p}_2$	$5+\sqrt{30}$	$\mathfrak{p}_4^*$	$13+2\sqrt{30}$
35	2	$6+\sqrt{35}$	1	$\mathfrak{p}_1$	3	$\mathfrak{p}_3$	$\mathfrak{p}_4$	–
42	2	$13+2\sqrt{42}$	1	$\mathfrak{p}_1$	$\mathfrak{p}_2$	5	$7+\sqrt{42}$	–
70	2	$251+30\sqrt{70}$	1	$\mathfrak{p}_1$	$\mathfrak{p}_2^*$	$25+3\sqrt{70}$	$\mathfrak{p}_4$	$17+2\sqrt{70}$
105	2	$41+4\sqrt{105}$	1	$\mathfrak{p}_1^*$	$\mathfrak{p}_2$	$10+\sqrt{105}$	$\mathfrak{p}_4$	$\frac{1}{2}(11+\sqrt{105})$
210	4	$29+2\sqrt{210}$	1	$\mathfrak{p}_1$	$\mathfrak{p}_2$	$\mathfrak{p}_3$	$\mathfrak{p}_4$	–

Table III.

D	$\mathfrak{p}_1\cdot\mathfrak{p}_2$	$\mathfrak{p}_1\cdot\mathfrak{p}_3$	$\mathfrak{p}_1\cdot\mathfrak{p}_4$	$\mathfrak{p}_2\cdot\mathfrak{p}_3$	$\mathfrak{p}_2\cdot\mathfrak{p}_4$	$\mathfrak{p}_3\cdot\mathfrak{p}_4$
10	$-2+\sqrt{10}$	$\sqrt{10}$	–	$5-\sqrt{10}$	–	–
15	$3+\sqrt{15}$	$5+\sqrt{15}$	$1+\sqrt{15}$	$\sqrt{15}$	$6-\sqrt{15}$	$-5+2\sqrt{15}$
30	$6+\sqrt{30}$	–	$-4+\sqrt{30}$	–	$3+\sqrt{30}$	–
35	–	$5+\sqrt{35}$	$7+\sqrt{35}$	–	–	$\sqrt{35}$
42	$6+\sqrt{42}$	–	–	–	–	–
70	$-8+\sqrt{70}$	–	$42+5\sqrt{70}$	–	$7+\sqrt{70}$	–
105	$\frac{1}{2}(-9+\sqrt{105})$	–	$\frac{1}{2}(7+\sqrt{105})$	–	$21+2\sqrt{105}$	–
210	–	–	$14+\sqrt{210}$	$15+\sqrt{210}$	–	–

Table IV.

D	α	I	I_U	D	α	I	I_U	D	α	I	I_U
2	1	–	2357	14	$4+\sqrt{14}$	–	7	35	1	–	2357
	1	7	235		$4+\sqrt{14}$	5	7		$\sqrt{35}$	–	23
	$\sqrt{2}$	–	3 7		$7+2\sqrt{14}$	–	2		$5+\sqrt{35}$	–	7
	$\sqrt{2}$	7	35		$7+2\sqrt{14}$	5	2		$7+\sqrt{35}$	–	5
3	1	–	2357	15	1	–	2357	42	1	–	2357
	$\sqrt{3}$	–	2 7		1	7	235		$\sqrt{42}$	–	–
	$1+\sqrt{3}$	–	3		$\sqrt{15}$	–	2		$6+\sqrt{42}$	–	57
	$3+\sqrt{3}$	–	5		$\sqrt{15}$	7	2		$7+\sqrt{42}$	–	3
5	2	–	2357		$3+\sqrt{15}$	–	57	70	1	–	2357
	$2\sqrt{5}$	–	23 7		$3+\sqrt{15}$	7	5		1	3	2 57
6	1	–	2357		$5+\sqrt{15}$	–	3		$\sqrt{70}$	–	–
	1	5	23 7		$5+\sqrt{15}$	7	3		$\sqrt{70}$	3	–
	$\sqrt{6}$	–	57		$1+\sqrt{15}^*$	7	35		$25+3\sqrt{70}$	–	3 7
	$\sqrt{6}$	5	7		$15+\sqrt{15}^*$	7	–		$25+3\sqrt{70}$	3	7
	$2+\sqrt{6}$	–	3		$6-\sqrt{15}^*$	7	2 5		$42+5\sqrt{70}$	–	5
	$2+\sqrt{6}$	5	3		$-5+2\sqrt{15}^*$	7	23		$42+5\sqrt{70}$	3	5
	$3+\sqrt{6}$	–	–	21	2	–	2357		$7+\sqrt{70}^*$	3	5
	$3+\sqrt{6}$	5	2		2	5	23 7		$10+\sqrt{70}^*$	3	7
7	1	–	2357		$2\sqrt{21}$	–	2 5		$-8+\sqrt{70}^*$	3	57
	1	3	2 57		$2\sqrt{21}$	5	2		$35-4\sqrt{70}^*$	3	2
	$\sqrt{7}$	–	2		$3+\sqrt{21}$	–	2 7	105	2	–	2357
	$\sqrt{7}$	3	2 5		$3+\sqrt{21}$	5	2 7		2	2	357
	$3+\sqrt{7}$	–	7		$7+\sqrt{21}$	–	23		$2\sqrt{105}$	–	2
	$3+\sqrt{7}$	3	57		$7+\sqrt{21}$	5	23		$2\sqrt{105}$	2	–
	$7+3\sqrt{7}$	–	35	30	1	–	2357		$20+2\sqrt{105}$	–	23 7
	$7+3\sqrt{7}$	3	5		1	7	235		$20+2\sqrt{105}$	2	3 7
10	1	–	2357		$\sqrt{30}$	–	–		$42+4\sqrt{105}$	–	2 5
	1	3	2 57		$\sqrt{30}$	7	–		$42+4\sqrt{105}$	2	5
	$\sqrt{10}$	–	3 7		$5+\sqrt{30}$	–	3 7		$7+\sqrt{105}^*$	2	35
	$\sqrt{10}$	3	7		$5+\sqrt{30}$	7	3		$15+\sqrt{105}^*$	2	7
	$-2+\sqrt{10}^*$	3	57		$6+\sqrt{30}$	–	5		$-9+\sqrt{105}^*$	2	57
	$5-\sqrt{10}^*$	3	2 7		$6+\sqrt{30}$	7	5		$35-3\sqrt{105}^*$	2	3
14	1	–	2357		$3+\sqrt{30}^*$	7	5	210	1	–	2357
	1	5	23 7		$10+\sqrt{30}^*$	7	3		$\sqrt{210}$	–	–
	$\sqrt{14}$	–	35		$-4+\sqrt{30}^*$	7	35		$14+\sqrt{210}$	–	35
	$\sqrt{14}$	5	3		$15-2\sqrt{30}^*$	7	2		$15+\sqrt{210}$	–	7

Table V.

						p = 2							p = 3							p = 5							p = 7						
D	A	B	a	b	n_0	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7
2	2	-1	1	0	0	2	0	8	3	1	0	0	1	0	12	2	2	1	1	1	0	15	0	0	2	0	1	0	42	1	0	1	2
2	2	-1	0	1	0	0							1	6	12	0	2	0	0	0							1	21	42	0	0	0	2
3	4	1	1	0	0	3	0	8	4	0	0	1	1	0	9	0	2	1	0	1	0	15	0	1	2	0	1	0	28	3	0	0	2
3	4	1	0	1	0	1							0							0							1	14	28	0	0	0	2
3	4	1	1	1	-1	0							1	4	9	0	2	0	0	0							0						
3	4	1	3	1	-1	0							0							1	7	15	0	0	2	0	0						
5	1	-1	2	0	0	4	0	24	5	2	0	1	2	0	36	4	3	0	0	1	0	25	0	0	2	0	1	0	56	0	1	0	2
5	1	-1	0	2	0	2							2	18	36	1	3	0	0	0							1	28	56	0	0	0	2
6	10	1	1	0	0	3	0	8	4	0	1	2	0	0	3	1	2	0	0	1	0	10	2	0	2	0	2	0	28	3	0	1	3
6	10	1	0	1	0	0							0							1	5	10	0	0	2	0	2	14	28	0	0	0	3
6	10	1	2	1	-1	0							2	4	9	0	3	0	0	0							0						
6	10	1	3	1	-1	0							0							0							0						
7	16	1	1	0	0	4	0	4	5	1	0	0	1	0	3	0	2	1	0	0	0	3	0	2	1	0	0	0	7	0	1	0	1
7	16	1	0	1	0	3							0							0							0						
7	16	1	3	1	-1	0							0							0							0	3	7	0	0	0	1
7	16	1	7	3	-1	0							2	4	9	0	3	1	0	1	7	15	0	2	2	0	0						
10	6	-1	1	0	0	1	0	4	2	1	0	0	1	0	6	1	2	0	0	0	0	5	0	0	1	0	0	0	8	3	1	0	1
10	6	-1	0	1	0	0							1	3	6	0	2	0	0	0							0	4	8	0	0	0	1
14	30	1	1	0	0	3	0	4	4	1	1	0	1	0	6	3	2	1	0	1	0	10	3	1	2	0	0	0	7	2	0	0	1
14	30	1	0	1	0	0							1	3	6	0	2	1	0	1	5	10	0	1	2	0	0						
14	30	1	4	1	-1	0							0							0							0	3	7	0	0	0	1
14	30	1	7	2	-1	1							0							0							0						
15	8	1	1	0	0	3	0	4	4	0	0	0	2	0	9	0	3	0	1	0	0	5	0	0	1	0	1	0	21	0	2	0	2
15	8	1	0	1	0	2							0							0							0						
15	8	1	3	1	-1	0							0							0	2	5	0	0	1	0	1	10	21	0	0	0	2
15	8	1	5	1	-1	0							2	4	9	0	3	0	0	0							0						
21	5	1	2	0	0	3	0	6	4	1	1	0	1	0	9	3	2	0	0	1	0	10	0	0	2	0	0	0	7	0	0	0	1
21	5	1	0	2	0	1							0							1	5	10	0	0	2	0	0						
21	5	1	3	1	-1	2							0							0							0	3	7	0	0	0	1
21	5	1	7	1	-1	1							1	4	9	1	2	0	0	0							0						
30	22	1	1	0	0	1	0	2	2	0	0	0	0	0	3	1	1	0	1	0	0	5	1	0	1	0	0	0	3	1	1	0	1
30	22	1	0	1	0	0							0							0							0						
30	22	1	5	1	-1	0							1	4	9	0	2	0	1	0							1	10	21	0	1	0	2
30	22	1	6	1	-1	0							0							0	2	5	0	0	1	0	0						
35	12	1	1	0	0	2	0	4	3	1	0	0	1	0	6	2	2	0	0	0	0	5	0	0	1	0	0	0	7	0	0	0	1
35	12	1	0	1	0	1							1	3	6	1	2	0	0	0							0						
35	12	1	5	1	-1	0							0							0							0	3	7	0	0	0	1
35	12	1	7	1	-1	0							0							0	2	5	0	0	1	0	0						

Table V. (cont.)

						p = 2							p = 3							p = 5							p = 7						
D	A	B	a	b	n_0	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7	ℓ_1	n_1	a_1	h_2	h_3	h_5	h_7
42	26	1	1	0	0	1	0	2	2	0	0	0	3	0	9	1	4	2	0	2	0	15	1	3	3	0	0	0	7	1	0	0	1
42	26	1	0	1	0	0							0							0							0						
42	26	1	6	1	-1	0							0							2	7	15	0	0	3	0	0	3	7	0	0	0	1
42	26	1	7	1	-1	0							3	4	9	0	4	0	0	0							0						
70	502	1	1	0	0	1	0	2	2	1	1	0	1	0	3	1	2	1	0	1	0	5	1	1	2	0	0	0	7	1	1	1	1
70	502	1	0	1	0	0							0							0							0						
70	502	1	25	3	-1	0							1	1	3	0	2	0	0	0							0	3	7	0	1	0	1
70	502	1	42	5	-1	0							0							1	2	5	0	0	2	0	0						
105	82	1	2	0	0	3	0	2	4	0	0	0	0	0	3	3	4	0	0	0	0	5	3	0	1	0	0	0	7	3	0	0	1
105	82	1	0	2	0	1							0							0							0						
105	82	1	20	2	-1	1							4	4	9	1	5	0	0	0							0	3	7	1	0	0	1
105	82	1	42	4	-1	2							0							0	2	5	2	0	1	0	0						
210	58	1	1	0	0	1	0	2	2	0	0	0	0	0	3	1	1	0	0	0	0	5	1	0	1	0	0	0	7	1	0	0	1
210	58	1	0	1	0	0							0							0							0						
210	58	1	14	1	-1	0							0	1	3	0	1	0	0	0	2	5	0	0	1	0	0						
210	58	1	15	1	-1	0							0							0							0	3	7	0	0	0	1

$(\alpha = a + b\sqrt{D})$

Table VI.

D	p_i	ν_i	λ_i	($i \in I_U^*$)
2	2 3 5	3 0 0	1.5 0 0	
6	2 3 7	3 1 0	1.5 0.5 0	
7	2 5 7	2 0 1	1 0 0.5	
10	2 5 7	3 1 0	1.5 0.5 0	
14	2 3 7	3 0 1	1.5 0 0.5	
15	2 3 5	2 1 1	1 0.5 0.5	
21	2 3 7	2 1 1	0 0.5 0.5	
30	2 3 5	3 1 1	1.5 0.5 0.5	
70	2 5 7	3 1 1	1.5 0.5 0.5	
105	3 5 7	1 1 1	0.5 0.5 0.5	

D	α	n_ε	n_π	n_2	n_3	n_5	n_7	I_U	I_U^*	N	κ	C_{12}^*
2	1	0	0	0	0	0	0	2 3 5	2 3 5	3	0	3.190×10^{28}
	$\sqrt{2}$	0	0	1	0	0	0	3 5	2 3 5	2	0	3.190×10^{28}
6	1	0	0	0	0	0	0	2 3 7	2 3 7	3	0	2.712×10^{26}
	$\sqrt{6}$	0	0	1	1	0	0	7	2 7	2	0	4.604×10^{22}
	$2+\sqrt{6}$	1	0	1	0	0	0	3	2 3	2	0	2.090×10^{22}
	$3+\sqrt{6}$	1	0	0	1	0	0	2	2 3	3	0	2.090×10^{22}
7	1	0	0	0	0	0	0	2 5 7	2 5 7	2	0	1.065×10^{30}
	$\sqrt{7}$	0	0	0	0	0	1	2 5	2 5	2	0	2.146×10^{28}
	$3+\sqrt{7}$	1	0	1	0	0	0	5 7	2 5 7	1	0	1.065×10^{30}
	$7+3\sqrt{7}$	1	0	1	0	0	1	5	2 5	1	0	2.146×10^{25}
10	1	0	0	0	0	0	0	2 5 7	2 5 7	3	0	3.214×10^{29}
	$\sqrt{10}$	0	0	1	0	1	0	7	2 7	2	0	8.414×10^{24}
	$-2+\sqrt{10}$	-1	1	1	0	0	0	5 7	2 5 7	2	1	3.214×10^{29}
	$5-\sqrt{10}$	-1	1	0	0	1	0	2 7	2 7	3	1	8.414×10^{24}
14	1	0	0	0	0	0	0	2 3 7	2 3 7	3	0	4.791×10^{26}
	$\sqrt{14}$	0	0	1	0	0	1	3	2 3	2	0	4.347×10^{22}
	$4+\sqrt{14}$	1	0	1	0	0	0	7	2 7	2	0	8.143×10^{22}
	$7+2\sqrt{14}$	1	0	0	0	0	1	2	2	3	0	8.371×10^{18}

Table VI. (cont.)

D	α	n_ε	n_π	n_2	n_3	n_5	n_7	I_U	I_U^*	N	κ	C_{12}^*
15	1	0	0	0	0	0	0	2 3 5	2 3 5	2	0	2.144×10^{28}
	$\sqrt{15}$	0	0	0	1	1	0	2	2	2	0	9.427×10^{19}
	$3+\sqrt{15}$	1	0	1	1	0	0	5	2 5	1	0	1.694×10^{24}
	$5+\sqrt{15}$	1	0	1	0	1	0	3	2 3	1	0	1.035×10^{24}
	$1+\sqrt{15}$	0	1	1	0	0	0	3 5	2 3 5	1	1	2.144×10^{28}
	$15+\sqrt{15}$	0	1	1	1	1	0		2	1	1	9.427×10^{19}
	$6-\sqrt{15}$	-1	1	0	1	0	0	2 5	2 5	2	1	1.694×10^{24}
	$-5+2\sqrt{15}$	-1	1	0	0	1	0	2 3	2 3	2	1	1.035×10^{24}
21	2	0	0	2	0	0	0	2 3 7	2 3 7	1	0	1.898×10^{26}
	$2\sqrt{21}$	0	0	2	1	0	1	2	2	0	0	2.640×10^{18}
	$3+\sqrt{21}$	1	0	2	1	0	0	2 7	2 7	1	0	3.220×10^{22}
	$7+\sqrt{21}$	1	0	2	0	0	1	2 3	2 3	1	0	1.435×10^{22}
30	1	0	0	0	0	0	0	2 3 5	2 3 5	3	0	4.141×10^{28}
	$\sqrt{30}$	0	0	1	1	1	0		2	2	0	2.022×10^{20}
	$5+\sqrt{30}$	1	0	0	0	1	0	3	2 3	3	0	2.217×10^{24}
	$6+\sqrt{30}$	1	0	1	1	0	0	5	2 5	2	0	3.276×10^{24}
	$3+\sqrt{30}$	0	1	0	1	0	0	5	2 5	3	1	3.276×10^{24}
	$10+\sqrt{30}$	0	1	1	0	1	0	3	2 3	2	1	2.217×10^{24}
	$-4+\sqrt{30}$	-1	1	1	0	0	0	3 5	2 3 5	2	1	4.141×10^{28}
	$15-2\sqrt{30}$	-1	1	0	1	1	0	2	2	3	1	2.022×10^{20}
70	1	0	0	0	0	0	0	2 5 7	2 5 7	3	0	3.229×10^{30}
	$\sqrt{70}$	0	0	1	0	1	1		2	2	0	2.115×10^{21}
	$25+3\sqrt{70}$	1	0	0	0	1	0	7	2 7	3	0	8.482×10^{25}
	$42+5\sqrt{70}$	1	0	1	0	0	1	5	2 5	2	0	7.003×10^{25}
	$7+\sqrt{70}$	0	1	0	0	0	1	5	2 5	3	1	7.003×10^{25}
	$10+\sqrt{70}$	0	1	1	0	1	0	7	2 7	2	1	8.482×10^{25}
	$-8+\sqrt{70}$	-1	1	1	0	0	0	5 7	2 5 7	2	1	3.229×10^{30}
	$35-4\sqrt{70}$	-1	1	0	0	1	1	2	2	3	1	2.115×10^{21}
105	2	0	0	2	0	0	0	3 5 7	3 5 7	1	0	4.533×10^{29}
	$2\sqrt{105}$	0	0	2	1	1	1			0	0	4.295×10^{16}
	$20+2\sqrt{105}$	1	0	2	0	1	0	3 7	3 7	1	0	1.690×10^{25}
	$42+4\sqrt{105}$	1	0	2	1	0	1	5	5	1	0	8.655×10^{20}
	$7+\sqrt{105}$	0	1	2	0	0	1	3 5	3 5	1	1	1.396×10^{25}
	$15+\sqrt{105}$	0	1	2	1	1	0	7	7	1	1	1.049×10^{21}
	$-9+\sqrt{105}$	-1	1	2	1	0	0	5 7	5 7	1	1	2.485×10^{25}
	$35-3\sqrt{105}$	-1	1	2	0	1	1	3	3	1	1	5.880×10^{20}

Chapter 8. The Thue equation.

Acknowledgements. The research for this chapter has been done in cooperation with N. Tzanakis from Iraklion. The results have been published in Tzanakis and de Weger [1989[a]].

8.1. Introduction.

Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a binary form with integral coefficients, of degree at least three, and irreducible. Let m be a nonzero integer. The diophantine equation

$$F(X,Y) = m$$

in $X, Y \in \mathbb{Z}$ is called a *Thue equation*. It plays a central role in the theory of diophantine equations. In 1909 Thue proved that it has only finitely many solutions (cf. Thue [1909]). His proof was ineffective. An effective proof was given by Baker [1968]. See Chapter 5 of Shorey and Tijdeman [1986] for a survey of results on Thue equations. By using Lemma 2.4 in Baker's argument, we derive a fully explicit upper bound for the solutions of the Thue equation. Then we show how the methods developed in Chapter 3 can be used to actually find all the solutions of a Thue equation. Our method works in principle for any Thue equation, and in practice for any Thue equation of not too large degree, provided that some algebraic data on the form F are available. See also Tzanakis [1989] for a short introduction.

Variants of the method we use here have been used in practice to solve Thue equations by Ellison, Ellison, Pesek, Stahl and Stall [1975], Steiner [1986], Pethö and Schulenberg [1987], and Blass, Glass, Meronk and Steiner [1987[a]], [1987[b]]. In all these cases $m = 1$, whereas de Weger [1989[b]] treats an example with $m > 1$, using the method described in this chapter. When determining all cubes in the Fibonacci sequence, Pethö [1983] solved a Thue equation by the Gelfond-Baker method, but with a completely different way to find all the solutions below the upper bound. And there are numerous Thue equations that have been solved by different (usually ad hoc) methods.

8.2. From the Thue equation to a linear form in logarithms.

In this section we show how the general Thue equation leads to an inequality involving a linear form in the logarithms of algebraic numbers with rational integral coefficients (unknowns). Let

$$F(X,Y) = \sum_{i=0}^{n} f_i \cdot X^{n-i} \cdot Y^i \in \mathbb{Z}[X,Y]$$

be a binary form of degree $n \geqslant 3$ and let m be a nonzero integer. Consider the Thue equation

$$F(X,Y) = m \ , \tag{8.1}$$

in the unknowns $X,\ Y \in \mathbb{Z}$. If F is reducible over $\mathbb{Q}$, then (8.1) can be reduced to a system of finitely many equations of type (8.1) with irreducible binary forms. For such equations of degree 1 or 2 it is well known how to determine the solutions. Therefore we may assume from now on that F is irreducible over $\mathbb{Q}$ and of degree $\geqslant 3$. Let $g(x) = F(x,1)$. If $g(x) = 0$ has no real roots then one can trivially find small upper bounds for $\max(|X|,|Y|)$ for the solutions (X,Y) of (8.1). Therefore, throughout this chapter we assume that the algebraic equation $g(x) = 0$ has at least one real root. We number its roots as follows: $\xi^{(1)}, \ldots, \xi^{(s)}$ (with $s \geqslant 1$) are the real roots and $\xi^{(s+1)} = \overline{\xi^{(s+t+1)}}, \ldots, \xi^{(s+t)} = \overline{\xi^{(s+2t)}}$ are the non-real roots, so that we have t ($\geqslant 0$) pairs of complex-conjugate roots, and $s + 2 \cdot t = n$.

Consider the field $K = \mathbb{Q}(\xi)$, where $g(\xi) = 0$. We will define three positive real numbers $Y_1 < Y_2 < Y_3$, that will divide the set of possible solutions (X,Y) of (8.1) into four classes:

→ the 'very small' solutions, with $|Y| \leqslant Y_1$. They will be found by enumeration of all possibilities,

→ the 'small' solutions, with $Y_1 < |Y| \leqslant Y_2$. They will be found by evaluating the continued fraction expansions of the real roots $\xi^{(i)}$.

→ the 'large' solutions, with $Y_2 < |Y| \leqslant Y_3$. They will be proved not to exist by a computational diophantine approximation technique,

→ the 'very large' solutions, with $|Y| > Y_3$. They will be proved not to exist by the theory of linear forms in logarithms.

The value of Y_3 follows from the Gelfond-Baker theory of linear forms in logarithms. The value of Y_2 follows from the restrictions that we use as we try to prove that no 'large' solutions exist. The value of Y_1 follows from Lemma 8.1 below. This lemma shows that if $|Y|$ is large enough then X/Y is 'extremely close' to one of the real roots $\xi^{(i)}$. In a typical example Y_3 may be as large as $10^{10^{50}}$, Y_2 as 10^{10}, and Y_1 as small as 10.

<u>LEMMA 8.1.</u> *Let* $X, Y \in \mathbb{Z}$ *satisfy* (8.1). *Put* $\beta = X - \xi \cdot Y$,

$$Y_0 = \begin{cases} \left\lceil \left(\dfrac{2^{n-1} \cdot |m|}{\min\limits_{1 \leqslant i \leqslant t} |g'(\xi^{(s+i)})| \cdot \min\limits_{1 \leqslant i \leqslant t} |\mathrm{Im}\, \xi^{(s+i)}|} \right)^{1/n} \right\rceil & \text{if } t \geqslant 1 \\ 1 & \text{if } t = 0 \end{cases} ,$$

$$C_1 = \frac{2^{n-1} \cdot |m|}{\min\limits_{1 \leqslant i \leqslant s} |g'(\xi^{(i)})|} , \quad C_2 = \tfrac{1}{2} \cdot \min_{1 \leqslant i < j \leqslant n} |\xi^{(i)} - \xi^{(j)}| ,$$

$$Y_1 = \max\left(Y_0, \left\lceil (4 \cdot C_1)^{1/(n-2)} \right\rceil \right) .$$

(i). *If* $|Y| > Y_0$ *then there exists an* $i_0 \in \{ 1, \ldots, s \}$ *such that*

$$|\beta^{(i_0)}| \leqslant C_1 \cdot |Y|^{-(n-1)} ,$$

$$|\beta^{(i)}| \geqslant C_2 \cdot |Y| \quad \text{for} \quad i \in \{ 1, \ldots, n \} , \; i \neq i_0 .$$

(ii). *If* $|Y| > Y_1$ *then* X/Y *is a convergent from the continued fraction expansion of* $\xi^{(i_0)}$.

<u>Proof.</u> Let $i_0 \in \{ 1, \ldots, n \}$ be such that $|\beta^{(i_0)}| = \min\limits_{1 \leqslant i \leqslant n} |\beta^{(i)}|$. We have from (8.1)

$$|f_0| \cdot \prod_{i=1}^{n} |\beta^{(i)}| = |m| .$$

By the minimality of $|\beta^{(i_0)}|$ we have for all i

$$|Y| \cdot |\xi^{(i)} - \xi^{(i_0)}| = |\beta^{(i)} - \beta^{(i_0)}| \leqslant |\beta^{(i)}| + |\beta^{(i_0)}| \leqslant 2 \cdot |\beta^{(i)}| .$$

Hence $|\beta^{(i)}| \geqslant C_2 \cdot |Y|$. Further,

$$|\beta^{(i_0)}| = \frac{|m|}{|f_0|} \cdot \prod_{i \neq i_0} |\beta^{(i)}|^{-1} \leqslant \frac{|m|}{|f_0|} \cdot \prod_{i \neq i_0} \left(\tfrac{1}{2} \cdot |Y| \cdot |\xi^{(i)} - \xi^{(i_0)}| \right)^{-1}$$

$$= \frac{2^{n-1}\cdot|m|}{\left|f_0\cdot\prod_{i\neq i_0}(\xi^{(i)}-\xi^{(i_0)})\right|\cdot|Y|^{n-1}} = \frac{2^{n-1}\cdot|m|}{\left|g'(\xi^{(i_0)})\right|\cdot|Y|^{n-1}} .$$

Now, if $i_0 > s$ (and hence $t \geqslant 1$) then, by the definition of Y_0 ,

$$\left|\frac{X}{Y} - \xi^{(i_0)}\right| = \frac{|\beta^{(i_0)}|}{|Y|} \leqslant \frac{2^{n-1}\cdot|m|}{\left|g'(\xi^{(i_0)})\right|}\cdot|Y|^{-n}$$

$$\leqslant \left(\frac{Y_0}{|Y|}\right)^n\cdot \min_{s+1\leqslant i\leqslant s+t} |\operatorname{Im}\xi^{(i)}| ,$$

which is impossible if $|Y| > Y_0$. Hence $i_0 \leqslant s$, and now (i) follows at once. Moreover, if $|Y| > Y_1$, then

$$\left|\frac{X}{Y} - \xi^{(i_0)}\right| = |\beta^{(i_0)}|\cdot|Y|^{-1} \leqslant C_1\cdot|Y|^{-n} \leqslant \tfrac{1}{4}\cdot Y_1^{n-2}\cdot|Y|^{-n} \leqslant \tfrac{1}{2}\cdot|Y|^{-2} ,$$

and thus $\left| \frac{X}{Y} - \xi^{(i_0)}\right| < \frac{1}{2}\cdot|Y|^{-2}$, since $\xi^{(i_0)}$ is irrational. Now (ii) follows from a well known result on continued fractions, cf. (3.6). □

Now let $|Y| > Y_1$ and $i_0 \in \{ 1, \ldots, s \}$ as in Lemma 8.1. Choose $j, k \in \{ 1, \ldots, n \}$ such that i_0, j, k are pairwise distinct and either $j, k \in \{ 1, \ldots, s \}$ or $j + t = k$ (so that $\xi^{(k)} = \overline{\xi^{(j)}}$), but further the choice of j, k is free. By $\beta^{(i)} = X - Y\cdot\xi^{(i)}$ for $i = i_0, j, k$ we get, on eliminating the X and Y ,

$$\beta^{(i_0)}\cdot\left(\xi^{(j)}-\xi^{(k)}\right) + \beta^{(j)}\cdot\left(\xi^{(k)}-\xi^{(i_0)}\right) + \beta^{(k)}\cdot\left(\xi^{(i_0)}-\xi^{(j)}\right) = 0 ,$$

or, equivalently,

$$\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}} - 1 = -\frac{\xi^{(k)}-\xi^{(j)}}{\xi^{(k)}-\xi^{(i_0)}}\cdot\frac{\beta^{(i_0)}}{\beta^{(j)}} . \tag{8.2}$$

By Lemma 8.1, the right hand side of (8.2) is 'extremely small'. Put, if $j, k \in \{1, \ldots, s \}$ (let us call it 'the real case')

$$\Lambda = \log\left|\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\cdot\frac{\beta^{(k)}}{\beta^{(j)}}\right|$$

and if $j, k \in \{ s+1, \ldots, s+2\cdot t \}$ (let us call it 'the complex case')

$$\Lambda = \frac{1}{i}\cdot \mathrm{Log}\left(\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}} \cdot \frac{\beta^{(k)}}{\beta^{(j)}} \right) ,$$

where, in general, for $z \in \mathbb{C}$, $\mathrm{Log}(z)$ denotes the principal value of the logarithm of z (hence $-\pi < \mathrm{Im}\,\mathrm{Log}(z) \leqslant \pi$). By $\xi^{(k)} = \overline{\xi^{(j)}}$ we have $\Lambda \in \mathbb{R}$ and $|\Lambda| \leqslant \pi$.

The following lemma shows how small $|\Lambda|$ is.

LEMMA 8.2. *Put*

$$C_3 = \max_{i_1 \neq i_2 \neq i_3 \neq i_1} \left| \frac{\xi^{(i_1)} - \xi^{(i_2)}}{\xi^{(i_1)} - \xi^{(i_3)}} \right| ,$$

$$Y_2^* = \max \left(Y_1, \left\lceil (2\cdot C_1 \cdot C_3/C_2)^{1/n} \right\rceil \right) .$$

If $|Y| > Y_2^*$ *then*

$$|\Lambda| < \frac{1.39\cdot C_1 \cdot C_3}{C_2}\cdot |Y|^{-n} .$$

Proof. Consider first the real case. From $|Y| > Y_2^*$ and Lemma 8.1 it follows that the right hand side of (8.2) is absolutely less than $\frac{1}{2}$ and, consequently,

$$\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}} \cdot \frac{\beta^{(k)}}{\beta^{(j)}} > 0 .$$

It follows that the left hand side of (8.2) is equal to $e^{\Lambda} - 1$, and now (8.2) implies, in view of Lemma 8.1 and the definition of C_3 ,

$$|e^{\Lambda}-1| < C_3 \cdot \frac{C_1 \cdot |Y|^{-(n-1)}}{C_2 \cdot |Y|} = \frac{C_1 \cdot C_3}{C_2}\cdot |Y|^{-n} .$$

On the other hand, $|e^{\Lambda}-1| < \frac{1}{2}$ implies (cf. Lemma 2.2)

$$|\Lambda| \leqslant 2\cdot \log 2 \cdot |e^{\Lambda}-1| \leqslant 1.39\cdot |e^{\Lambda}-1| ,$$

which proves our claim in the real case.
In the complex case the left hand side of (8.2) is equal to $e^{i\Lambda} - 1$, and, as in the real case, we derive

$$|e^{i\Lambda}-1| < \frac{C_1 \cdot C_3}{C_2} \cdot |Y|^{-n} < \tfrac{1}{2} .$$

Since $|e^{i\Lambda}-1| = 2 \cdot |\sin \Lambda/2|$, it follows that $|\sin \Lambda/2| < \frac{1}{4}$, and therefore by Lemma 2.3

$$|\Lambda| \leqslant 2 \cdot \frac{1/4}{\sin 1/4} \cdot |\sin \Lambda/2| = \frac{1/4}{\sin 1/4} \cdot |e^{i\Lambda}-1| \leqslant 1.02 \cdot |e^{i\Lambda}-1| ,$$

which proves the lemma in the complex case. □

In the ring of integers of the field K (as well as in any other order R of K) there exists a system of fundamental units $\varepsilon_1, \ldots, \varepsilon_r$, where $r = s + t - 1$ (Dirichlet's Unit Theorem). Note that since F is irreducible and we have supposed $s > 0$, the only roots of unity belonging to K are ± 1 . We shall not discuss here the problem of finding such a system (for efficient methods see e.g. Berwick [1932], Billevič [1956], [1964], Pohst and Zassenhaus [1982], Buchmann [1985], [1986]). We simply assume that a system of fundamental units is known. On the other hand, there exist only finitely many non-associates $\mu_1, \ldots, \mu_\nu$ in K such that $f_0 \cdot N(\mu_i) = m$ for $i = 1, \ldots, \nu$ (we use $N(\cdot)$ to denote the norm of the extension $K/\mathbb{Q}$). We also assume that a complete set of such μ_i's is known. Let M be the set of all $\zeta \cdot \mu_i$, where ζ is a root of unity in K . (In the important case $|f_0| = |m| = 1$, it is clear that $M = \{ -1, 1 \}$). Then, for any integral solution (X,Y) of (8.1) there exist some $\mu \in M$ and $a_1, \ldots, a_r \in \mathbb{Z}$, such that

$$\beta = \mu \cdot \varepsilon_1^{a_1} \cdot \ldots \cdot \varepsilon_r^{a_r} .$$

Thus, the initial problem of solving (8.1) is reduced to that of finding all integral r-tuples $(a_1, \ldots, a_r)$ such that $\mu \cdot \varepsilon_1^{a_1} \cdot \ldots \cdot \varepsilon_r^{a_r}$ for some $\mu \in M$ be of the special shape $X - Y \cdot \xi$, with $X, Y \in \mathbb{Z}$. As we have seen, X and Y can be eliminated, so that we obtain (8.2). Thus the problem reduces to solving finitely many equations of the type

$$\frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \cdot \frac{\mu^{(k)}}{\mu^{(j)}} \cdot \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right)^{a_i} - 1 = - \frac{\xi^{(k)} - \xi^{(j)}}{\xi^{(k)} - \xi^{(i_0)}} \cdot \frac{\mu^{(i_0)}}{\mu^{(j)}} \cdot \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}} \right)^{a_i}$$

(the so-called 'unit equation'). In the real case we have

$$\Lambda = \log \left| \frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \cdot \frac{\mu^{(k)}}{\mu^{(j)}} \right| + \sum_{i=1}^{r} a_i \cdot \log \left| \frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right| , \tag{8.3}$$

and in the complex case

$$\Lambda = \operatorname{Arg} \left(\frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}} \cdot \frac{\mu^{(k)}}{\mu^{(j)}} \right) + \sum_{i=1}^{r} a_i \cdot \operatorname{Arg}\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right) + a_0 \cdot 2\pi , \tag{8.4}$$

with $a_0 \in \mathbb{Z}$, and $-\pi < \operatorname{Arg}(z) \leqslant \pi$ for every $z \in \mathbb{C}$. Note that Λ in the real case, and $i \cdot \Lambda$ in the complex case, is a linear form in (principal) logarithms of algebraic numbers, where the coefficients a_i are integers. The Gelfond-Baker theory provides an explicit lower bound for $|\Lambda|$ in terms of $\max|a_i|$. Using this in combination with Lemma 8.2 we can find an explicit upper bound for $\max|a_i|$. This is what we do in the next section.

8.3. Upper bounds.

Let $A = \max_{1 \leqslant i \leqslant r} |a_i|$. First we find an upper bound for A in terms of $|Y|$.

<u>LEMMA 8.3.</u> *Let* $I = \{ h_1, \ldots, h_r \} \subset \{ 1, \ldots, n \}$. *Put*

$$U_I = \left(\log|\varepsilon_\ell^{(h_i)}| \right)_{1 \leqslant i \leqslant r, 1 \leqslant \ell \leqslant r} ,$$

(where i *indicates a row and* ℓ *a column of the matrix),*

$$U_I^{-1} = (u_{i\ell}) , \quad N[U_I^{-1}] = \max_{1 \leqslant i \leqslant r} \sum_{\ell=1}^{r} |u_{i\ell}| .$$

Put also

$$\mu_- = \min_{\substack{1 \leqslant i \leqslant n \\ \mu \in M}} |\mu^{(i)}| , \quad \mu_+ = \max_{\substack{1 \leqslant i \leqslant n \\ \mu \in M}} |\mu^{(i)}| ,$$

$$C_4 = \frac{\frac{1}{2} + \max_{1 \leqslant i_1 < i_2 \leqslant n} |\xi^{(i_1)} - \xi^{(i_2)}|}{\mu_-} ,$$

$$C_5 = \min \left((n-1) \cdot \min_I N[U_I^{-1}], \max_I N[U_I^{-1}] \right) .$$

Then, for

$$|Y| > \max\left(Y_1,\ 2\cdot|m|^{1/n},\ \mu_+/C_2 \right) ,$$

we have

$$A < C_5\cdot\log\left(C_4\cdot|Y|\right) .$$

<u>Proof.</u> By $\beta = \mu\cdot\varepsilon_1^{a_1}\cdot\ldots\cdot\varepsilon_r^{a_r}$ we have

$$\begin{pmatrix} \log|\beta^{(h_1)}/\mu^{(h_1)}| \\ \vdots \\ \log|\beta^{(h_r)}/\mu^{(h_r)}| \end{pmatrix} = U_I\cdot\begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} . \tag{8.5}$$

On the other hand, for every $h \in \{ 1, \ldots, n \}$, using the end of the proof of Lemma 8.1,

$$\begin{aligned} |\beta^{(h)}| = |X-Y\cdot\xi^{(h)}| &\leqslant |X-Y\cdot\xi^{(i_0)}| + |Y|\cdot|\xi^{(i_0)}-\xi^{(h)}| \\ &\leqslant \frac{1}{2\cdot|Y|} + |Y|\cdot|\xi^{(i_0)}-\xi^{(h)}| \\ &< \left(\tfrac{1}{2} + \max_{1\leqslant i_1<i_2\leqslant n}|\xi^{(i_1)}-\xi^{(i_2)}| \right)\cdot|Y| , \end{aligned}$$

and therefore

$$\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| < C_4\cdot|Y| \quad \text{for} \quad h = 1, \ldots, n .$$

Note that $C_4\cdot|Y| > 1$. Indeed, by

$$\prod_{i=1}^{n} |\mu^{(i)}| = \frac{|m|}{|f_0|} \leqslant |m|$$

it follows that $\min_{1\leqslant i\leqslant n} |\mu^{(i)}| \leqslant |m|^{1/n}$, hence $\mu_- \leqslant |m|^{1/n}$. Therefore

$$C_4\cdot|Y| \geqslant \left(\tfrac{1}{2} + \max_{1\leqslant i_1<i_2\leqslant n}|\xi^{(i_1)}-\xi^{(i_2)}| \right)\cdot|Y|\cdot|m|^{-1/n} > \frac{|Y|}{2|m|^{1/n}} > 1 .$$

Then,

$$\log\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| < \log\left(C_4\cdot|Y|\right) \text{ for } h = 1,\ldots,n ,\quad \log\left(C_4\cdot|Y|\right) > 0 . \tag{8.6}$$

Next we show that

$$\left|\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right|\right| < (n-1)\cdot\log\bigl(C_4\cdot|Y|\bigr) \quad \text{for} \quad i = 1, \ldots, n\ . \tag{8.7}$$

Indeed, in view of (8.6), a stronger inequality is true if $|\beta^{(i)}/\mu^{(i)}| \geqslant 1$. Suppose now that $|\beta^{(i)}/\mu^{(i)}| < 1$. By

$$\prod_{h=1}^{n}\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| = 1$$

it follows that

$$\left|\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right|\right| = -\log\left|\frac{\beta^{(i)}}{\mu^{(i)}}\right| = \sum_{\substack{h=1\\h\neq i}}^{n}\log\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right| < (n-1)\cdot\log\bigl(C_4\cdot|Y|\bigr)\ ,$$

in view of (8.6). Now the inequality

$$A < (n-1)\cdot\min_{I} N[U_I^{-1}]\cdot\log\bigl(C_4\cdot|Y|\bigr)$$

follows from (8.5), (8.7), the definition of $N[U_I^{-1}]$ and the fact that, as we have not put so far any restriction on I , this could be chosen so that $N[U_I^{-1}]$ be minimal. It remains to show that

$$A < \max_{I} N[U_I^{-1}]\cdot\log(C_4\cdot|Y|)\ .$$

Choose I such that $i_0 \notin I$. Then, by Lemma 8.1, for every $h \in I$, $|\beta^{(h)}/\mu^{(h)}| > C_2\cdot|Y|/\mu_+ > 1$ and now, in view of (8.6),

$$\left|\log\left|\frac{\beta^{(h)}}{\mu^{(h)}}\right|\right| < \log\bigl(C_4\cdot|Y|\bigr)\ ,$$

which implies our assertion. □

Lemmas 8.2 and 8.3 immediately yield

<u>LEMMA 8.4.</u> *Put*

$$C_6 = \frac{1.39\cdot C_1\cdot C_3\cdot C_4^n}{C_2}\ , \quad Y_2' = \max\bigl(\,Y_2^*,\ 2\cdot|m|^{1/n}\ ,\ \mu_+/C_2\,\bigr)\ .$$

If $|Y| > Y_2'$ *then*

$$|\Lambda| < C_6\cdot\exp\Bigl(\frac{-n}{C_5}\cdot A\Bigr)\ .$$

Next we apply Lemma 2.4 (Waldschmidt). It yields in the real case (assuming that $\Lambda \neq 0$)

$$|\Lambda| > \exp\bigl(-C_7\cdot(\log A + C_8)\bigr) , \tag{8.8}$$

and in the complex case this holds when A is replaced by $A' = \max\limits_{0\leqslant i\leqslant r} |a_i|$. The precise values for C_7 and C_8 are given in Section 2.3. It should be noted that in the complex case a_0 makes now its appearance, while it was not present in Lemmas 8.3 and 8.4. In order to obtain an upper bound for A , we must find an upper bound for A' in terms of A . Indeed, using

$$\mathrm{Arg}(z_1\cdot z_2) = \mathrm{Arg}(z_1) + \mathrm{Arg}(z_2) + k\cdot 2\pi , \quad k \in \{ -1, 0, 1 \} ,$$

we find from (8.4) and the proof of lemma 8.2 that if $A \geqslant 2$ then

$$|a_0| < \tfrac{1}{2} + \tfrac{1}{2}\cdot r\cdot A + |\Lambda|/2\pi < 1 + r\cdot A \leqslant r\cdot A .$$

Thus we may apply (8.8) in both cases with the same A if we replace C_8 by

$$C_8' = C_8 \qquad \text{in the real case,}$$

$$C_8' = C_8 + \log r \qquad \text{in the complex case.}$$

We can now give an upper bound for A .

LEMMA 8.5. *Put*

$$C_9 = \frac{2\cdot C_5}{n}\cdot\Bigl(\log C_6 + C_7\cdot C_8' + C_7\cdot\log\frac{C_5\cdot C_7}{n} \Bigr) .$$

If $|Y| > Y_2'$, *then* $A < C_9$.

Proof. As we have seen in the proof of Lemma 8.2, $|e^{\Lambda}-1| < \frac{1}{2}$ in the real case, and $|e^{i\Lambda}-1| < \frac{1}{2}$ in the complex case. Note that $\beta^{(i_0)} \neq 0$. Hence (8.2) implies $\Lambda \neq 0$. Therefore Lemma 8.4 and (8.8) yield

$$A < \frac{C_5}{n}\cdot\Bigl(\log C_6 + C_7\cdot C_8' + C_7\cdot\log A \Bigr) .$$

The result now follows from Lemma 2.1. □

Remark. From this upper bound for A an upper bound for $|Y|$ can be derived, thus a value for Y_3 (cf. Section 8.2). We shall not do this. Note that Y_2' (cf. Lemma 8.4) is not necessarily equal to Y_2 (cf. Section 8.2).

8.4. Reducing the upper bound.

We are now left with a problem of the following type. Let be given real numbers $\delta, \mu_1, \ldots, \mu_q$ ($q \geqslant 2$, the case $q = 1$ is trivial). Write

$$\Lambda = \delta + a_1 \cdot \mu_1 + \ldots + a_q \cdot \mu_q ,$$

where the a_i's belong to $\mathbb{Z}$, and put $A = \max_{1 \leqslant i \leqslant q} |a_i|$. If K_1, K_2, K_3 are given positive numbers, then find all q-tuples $(a_1, \ldots, a_q) \in \mathbb{Z}^q$ satisfying

$$|\Lambda| < K_1 \cdot \exp\left(-K_2 \cdot A\right) , \quad A < K_3 . \tag{8.9}$$

In our case, it follows from (8.3) or (8.4) how to define q, δ and the μ_i's , and from Lemmas 8.4 and 8.5 how to define K_1, K_2, K_3 . In general, K_1 and K_2 are 'small' constants, whereas K_3 is 'very large'. Put

$$\Lambda_0 = a_1 \cdot \mu_1 + \ldots + a_q \cdot \mu_q ,$$

so that $\Lambda = \delta + \Lambda_0$. We apply the methods of Chapter 3 to problem (8.9).

Below we distinguish three cases. In the first two we suppose that the μ_i's are $\mathbb{Q}$-independent.

(i). Let $\delta = 0$, so that $\Lambda = \Lambda_0$. Then the linear form is homogeneous, and we apply the method of Section 3.7.

(ii) Let $\delta \neq 0$. Then the linear form is inhomogeneous, and we apply the method of Section 3.8.

(iii). Suppose now that the μ_i's are $\mathbb{Q}$-dependent. Let Γ be the approximation lattice for the linear form Λ , as defined in Section 3.7. Then we expect the lower bound for $|\underline{x}|$ ($\underline{x} \in \Gamma$, $\underline{x} \neq \underline{0}$) in general to be 'very small', since the vector having as coordinates the coefficients of the dependence relation will give rise to a very short vector in the lattice. So the reduction process, as applied in the two previous cases, will not work. In such a case we work as follows. Let M be a maximal subset of $\{\mu_1, \ldots, \mu_q\}$ consisting of $\mathbb{Q}$-independent numbers. With an appropriate choice of subscripts we may assume that $M = \{ \mu_1, \ldots, \mu_p \}$, $p < q$. Then we can find integers $d > 0$ and d_{ij} for $1 \leqslant i \leqslant p < j \leqslant q$ such that

$$d \cdot \mu_j = \sum_{i=1}^{p} d_{ij} \cdot \mu_i \quad \text{for} \quad j = p+1, \ldots, q .$$

(These numbers d, d_{ij} can be found as coordinates of extremely short vectors in reduced bases). On the other hand, (8.9) is equivalent to

$$|\Lambda'| < K_1'\cdot\exp\left(-K_2\cdot A\right)\ , \quad A < K_3\ , \tag{8.10}$$

where $\Lambda' = d\cdot\Lambda$ and $K_1' = d\cdot K_1$. Now, with $\delta' = d\cdot\delta$ and

$$a_i' = d\cdot a_i + \sum_{j=p+1}^{q} d_{ij}\cdot a_j$$

we obtain

$$\Lambda' = \delta' + \sum_{i=1}^{p} a_i'\cdot\mu_i\ .$$

Put $D = \max\left(\ |d|,\ |d_{ij}|\ :\ 1 \leqslant i \leqslant p < j \leqslant q\ \right)$. Then

$$|a_i'| \leqslant (q-p+1)\cdot D\cdot A \quad \text{for} \quad i = 1,\ \ldots,\ p\ .$$

Therefore, put $A' = \max_{1\leqslant i\leqslant p} |a_i'|$, then $A' \leqslant (q-p+1)\cdot D\cdot A$, and (8.10) implies

$$|\Lambda'| < K_1'\cdot\exp\left(-K_2'\cdot A'\right)\ , \quad A' < K_3'\ , \tag{8.11}$$

where

$$\Lambda' = \delta' + a_1'\cdot\mu_1' + \ldots + a_p'\cdot\mu_p'\ , \quad K_1' = d\cdot K_1\ ,$$

$$K_2' = K_2/(q-1+p)\cdot D\ , \quad K_3' = (q-p+1)\cdot K_3\ .$$

Now, to solve (8.11) we apply the reduction process described in (i) or (ii), depending on whether $\delta' = 0$ or $\delta' \neq 0$, and maybe more than once, if necessary, until we find a very small upper bound for A' . After having found all solutions $(a_1',\ldots,a_p')$ of (8.11), we have a lower bound $L > 0$ for $|\Lambda'|$. It is reasonable to expect that L is not 'extremely small', because the integers a_1', ..., a_p' being 'small' in absolute value cannot make $|\Lambda'|$ 'extremely small'. Now combine $|\Lambda'| \geqslant L$ with the first inequality of (8.10) to get

$$A < \frac{1}{K_2}\cdot\log\left(\frac{K_1}{L}\right)\ .$$

Since L is not 'very small', as argued heuristically, the above upper bound for A is 'small'.

Returning now to the general case, we point out that if the reduced upper bound for A (found after some reduction steps as described above) is not small enough to admit enumeration of the remaining possibilities in a reasonable time, then it might be necessary, or at least advisable, to use the technique of Fincke and Pohst, cf. Section 3.6. However, when solving a Thue equation, and not only an inequality for a linear form in logarithms, it may be better to avoid this method, and to use continued fractions of the roots $\xi^{(i)}$. In practice we can search for the solutions (X,Y) of (8.1) satisfying $Y_1 < |Y| \leq C$ as follows, referring to Lemma 8.1. Here e.g. $C = Y_2$, and we can imagine C here as being a 'large' constant compared to Y_1, but not 'very large' (cf. the introduction of Y_1, Y_2 in Section 8.2).

Let $\tilde{\xi}$ be a rational approximation of $\xi^{(i_0)}$, such that

$$\left|\tilde{\xi}-\xi^{(i_0)}\right| < \frac{1}{6\cdot C^2} . \tag{8.12}$$

Since $|Y| > Y_1$, X/Y must be a convergent, p_k/q_k say, from the continued fraction expansion of $\xi^{(i_0)}$. Denote by a_0, a_1, a_2, ... the partial quotients in this expansion. First we claim that $a_{k+1} \geq 3$. Indeed, by (3.5)

$$\frac{1}{(a_{k+1}+2)\cdot|Y|^2} \leq \frac{1}{(a_{k+1}+2)\cdot q_k^2} < \left|\xi^{(i_0)} - \frac{p_k}{q_k}\right| = \left|\xi^{(i_0)} - \frac{X}{Y}\right| \leq \frac{C_1}{|Y|^n} .$$

If $a_{k+1} = 1$ or 2, then we would have $|Y|^{n-2} < 4\cdot C_1$, which is absurd, since $|Y| > Y_1 > (4\cdot C_1)^{1/(n-2)}$. Thus, $a_{k+1} \geq 3$, and by (3.5) we have

$$\left|\xi^{(i_0)} - \frac{p_k}{q_k}\right| < \frac{1}{a_{k+1}\cdot q_k^2} \leq \frac{1}{3\cdot q_k^2}.$$

Therefore,

$$\left|\tilde{\xi} - \frac{p_k}{q_k}\right| \leq \left|\tilde{\xi}-\xi^{(i_0)}\right| + \left|\xi^{(i_0)} - \frac{p_k}{q_k}\right| < \frac{1}{6\cdot C^2} + \frac{1}{3\cdot q_k^2} \leq \frac{1}{2\cdot q_k^2}$$

and this means that p_k/q_k is in fact a convergent from the continued fraction expansion of $\tilde{\xi}$ too. Moreover, in view of the inequalities

$$\frac{1}{(a_{k+1}+2)\cdot q_k^2} < \left|\xi^{(i_0)} - \frac{p_k}{q_k}\right| \leq \frac{C_1}{|Y|^n} \leq \frac{C_1}{|q_k|^n} ,$$

a_{k+1} must be sufficiently large compared to q_k, namely

$$a_{k+1} > \frac{|q_k|^{n-2}}{C_1} - 2 . \tag{8.13}$$

This inequality can be checked easily for all k such that $q_k \leqslant C$.

To sum up, we propose the following process for every real root $\xi^{(i_0)}$ for $i_0 = 1, \ldots, s$ (note that i_0 is a priori not known). (1) Compute a rational approximation $\tilde{\xi}$ of $\xi^{(i_0)}$ satisfying (8.12) (a truncation of its decimal expansion will do). (2) Expand $\tilde{\xi}$ into its continued fraction with partial quotients $a_0, a_1, a_2, \ldots, a_{k+1}$ and convergents p_i/q_i for all $i = 1, \ldots, k$ with $q_k \leqslant C < q_{k+1}$. (3) Test all these convergents for the conditions (8.13) and $F(p_i,q_i) = m$. Concerning this last test, note that if $X/Y = p_i/q_i$, then $X = Z\cdot p_i$, $Y = Z\cdot q_i$ for some $Z \in \mathbb{Z}$ with $Z^n \mid m$. This simple observation excludes in general most of the reducible quotients X/Y , and all of them if m is an n-th-powerfree integer.

Having tested for all solutions in the range $|Y| \leqslant C$ we may suppose that $|Y| > C$. For such solutions (X,Y) we can obtain a lower bound for the corresponding A as follows (the idea is due to A. Pethö, cf. also Section 1 of Blass, Glass, Meronk and Steiner [1987[b]]). For every $(i,j) \in \{1,\ldots,r\} \times \{1,\ldots,n\}$ let φ_{ij} be the number $+1$ or -1 for which $|\varepsilon_i^{(j)}|^{\nu_{ij}} \geqslant 1$, and put $E_j = \prod_{i=1}^{r} |\varepsilon_i^{(j)}|^{\varphi_{ij}}$. Then

$$|\beta^{(j)}| = |\mu^{(j)}|\cdot \prod_{i=1}^{r} |\varepsilon_i^{(j)}|^{a_i} \leqslant \mu_+\cdot E_j^A$$

and hence for any pair j_1, j_2 with $j_1 \neq j_2$ we have

$$|Y| = \frac{|\beta^{(j_1)}-\beta^{(j_2)}|}{|\xi^{(j_1)}-\xi^{(j_2)}|} \leqslant \mu_+\cdot\frac{E_{j_1}^A + E_{j_2}^A}{|\xi^{(j_1)}-\xi^{(j_2)}|} ,$$

and from this we can find a lower bound for A , if we know that $|Y| > C$. Of course, for an other pair j_1, j_2 we may find a different lower bound, and therefore we can take the larger one.

8.5. An application: triangular numbers that are a product of three consecutive numbers.

In this section we prove, as an application of the general theory described in the previous sections, the following result. The problem was posed by S.P.

Mohanty (cf. Mohanty [1988]; the proof in this paper is incorrect). The *n-th triangular number* is for $n \in \mathbb{N}$ defined as $T_n = \frac{1}{2} \cdot n \cdot (n+1)$.

<u>THEOREM 8.6.</u> *The only triangular numbers that are a product of three consecutive integers, are* $T_3 = 1 \cdot 2 \cdot 3$, $T_{15} = 4 \cdot 5 \cdot 6$, $T_{20} = 5 \cdot 6 \cdot 7$, $T_{44} = 9 \cdot 10 \cdot 11$, $T_{608} = 56 \cdot 57 \cdot 58$, $T_{22736} = 636 \cdot 637 \cdot 638$.

<u>Proof.</u> We have the diophantine equation $n \cdot (n+1) = 2 \cdot m \cdot (m+1) \cdot (m+2)$ in $n, m \in \mathbb{N}$. Put $x = 2 \cdot m + 2$, $y = 2 \cdot n + 1$. Then we are lead to the equation $y^2 = x^3 - 4 \cdot x + 1$ in $x, y \in \mathbb{N}$, with $x \geqslant 4$ even and $y \geqslant 3$ odd. Theorem 8.7 below now completes the proof. □

<u>THEOREM 8.7.</u> *The elliptic curve*

$$y^2 = x^3 - 4 \cdot x + 1 \tag{8.14}$$

has only the following 22 integral points:

$$(x, \pm y) = (-2,1),\ (-1,2),\ (0,1),\ (2,1),\ (3,4),\ (4,7),\ (10,31),$$

$$(12,41),\ (20,89),\ (114,1217),\ (1274,45473)\ .$$

We prove this theorem in two main steps. First, we reduce the problem to the solution of two quartic Thue equations. Then we solve these equations using the general theory developed in the previous sections.

Let L be the totally real field $\mathbb{Q}(\psi)$, where

$$\psi^3 - 4 \cdot \psi + 1 = 0\ .$$

Let the conjugates of ψ be $\psi^{(1)} = 0.254\ldots$, $\psi^{(2)} = -2.114\ldots$, $\psi^{(3)} = 1.860\ldots$. From a table of Delone and Faddeev ([1964], p. 141) we see that the class number of L is 1 , its ring of integers is $\mathbb{Z}[\psi]$, its discriminant is 229 , and a pair of independent units is $\psi,\ 2 - \psi$. From Table I of Buchmann [1986] we see that $-7 + 2 \cdot \psi^2,\ 2 \cdot \psi + \psi^2$ is a pair of fundamental units in $\mathbb{Z}[\psi]$. By $-7 + 2 \cdot \psi^2 = -\psi^{-1} \cdot (2-\psi)$, $2 \cdot \psi + \psi^2 = (2-\psi)^{-1}$ we see that $\psi,\ 2 - \psi$ is also a pair of fundamental units in $\mathbb{Z}[\psi]$.

The equation (8.14) of the elliptic curve can be written as

$$y^2 = (x - \psi) \cdot (x^2 + x \cdot \psi + (\psi^2 - 4)) \tag{8.15}$$

and the factors on the right hand side are relatively prime. Indeed, if π were a common prime divisor of them, then π would divide

$$(x^2 + x\cdot\psi + (\psi^2-4)) - (x + 2\cdot\psi)\cdot(x - \psi) = 3\cdot\psi^2 - 4 ,$$

which is prime, since its norm is -229. Therefore we would have that π is a unit times this prime, and then by (8.15), $x - \psi = \text{unit}\times(3\cdot\psi^2-4)\times\text{square}$. Take norms, then we get $y^2 = \pm 229\times\text{square}$, which is clearly impossible.

Now (8.15) implies

$$x - \psi = \pm\psi^i\cdot(2-\psi)^j\cdot\alpha^2 , \quad \alpha \in \mathbb{Z}[\psi] , \quad i,\ j \in \{ 0,\ 1 \} . \tag{8.16}$$

Since (8.14) is trivial to solve for $x \leqslant 0$ (the only solutions with $x \leqslant 0$ are the first three pairs stated in the theorem), we may assume that $x \geqslant 1$. Since $\psi^{(1)} = 0.254...$, we see that the minus sign in (8.16) is impossible. Then, by $\psi^{(2)} = -2.114...$, $i \neq 1$. We conclude therefore that

$$x - \psi = (2-\psi)^j\cdot(u+v\cdot\psi+w\cdot\psi^2)^2 , \quad u,\ v,\ w \in \mathbb{Z} ,\ j \in \{ 0,\ 1 \} . \tag{8.17}$$

<u>First case: $j = 0$.</u> Then (8.17) implies, on equating corresponding coefficients in both sides,

$$x = u^2-2\cdot v\cdot w,\ w^2-2\cdot u\cdot v-8\cdot v\cdot w = 1,\ v^2+4\cdot w^2+2\cdot u\cdot w = 0 . \tag{8.18}$$

Note that w is odd and v is even, hence $4 \mid 2\cdot u\cdot w$, so u is even. Put $u = 2\cdot u_1$, $v = 2\cdot v_1$. The last equation of (8.18) now reads

$$w^2 + u_1\cdot w + v_1^2 = 0 .$$

Consider this as a quadratic equation in w. Its discriminant must be a square, z^2 say. Then

$$u_1^2 - 4\cdot v_1^2 = z^2 , \quad w = \tfrac{1}{2} (-u_1 \pm z) .$$

Note that u_1 and z have the same parity. We may assume $u \geqslant 0$.

First suppose that u_1 and z are even. Since $w^2 + u_1\cdot w + v_1^2 = 0$ and w is odd, we find $u_1 \equiv 2 \pmod 4$, and v_1 is odd. Put $u_1 = 2\cdot u_2$, $z = 2\cdot z_1$. Then $u_2^2 - v_1^2 = z_1^2$, where u_2 and v_1 are odd. By $u_2 \geqslant 0$ there exist $m,\ n \in \mathbb{Z}$ such that

$$u_2 = m^2 + n^2 \ , \quad v_1 = m^2 - n^2 \ , \quad z_1 = 2\cdot m\cdot n \ .$$

It follows that

$$u = 4\cdot(m^2+n^2) \ , \quad v = 2\cdot(m^2-n^2) \ , \quad w = -(m\pm n)^2 \ .$$

Since the sign of z, and thus that of n, is of no importance, we may assume $w = -(m+n)^2$. After substitution in the second equation of (8.18) we obtain the Thue equation

$$m^4 + 36\cdot m^3\cdot n + 6\cdot m^2\cdot n^2 - 28\cdot m\cdot n^3 + n^4 = 1 \ .$$

The left hand side can be factored as

$$(\ m + n\ )\cdot(\ m^3 + 35\cdot m^2\cdot n - 29\cdot m\cdot n^2 + n^3\ ) \ ,$$

and therefore it can be solved very easily. Its only solutions are $\pm(m,n) = (1,0), (0,1)$. They lead to $\pm(u,v,w) = (4,2,-1), (4,-2,-1)$, and then by (8.18) we find $x = 20, 12$ respectively, which furnish the solutions $(x,\pm y) = (20,89), (12,41)$ for (8.14).

Secondly, we suppose that u_1 and z are odd. Then v_1 is even, so by $u_1 \geqslant 0$ there exist $m, n \in \mathbb{Z}$ with

$$u_1 = m^2 + n^2 \ , \quad 2\cdot v_1 = 2\cdot m\cdot n \ , \quad z = m^2 - n^2 \ .$$

It follows that

$$u = 2\cdot(m^2+n^2) \ , \quad v = 2\cdot m\cdot n \ , \quad w = -m^2 \ \text{ or } \ w = -n^2 \ .$$

We may assume that $w = -m^2$. Substituting this in the second equation of (8.18) we find the Thue equation

$$m^4 + 8\cdot m^3\cdot n - 8\cdot m\cdot n^3 = 1 \ .$$

The left hand side is again reducible. The only solutions, as is easily seen, are $\pm(m,n) = (1,0), (1,1), (1,-1)$. Since m and n cannot have the same parity, only the first pair is accepted. It leads to $(u,v,w) = (2,0,-1)$, and hence to $(x,\pm y) = (4,7)$ for (8.14).

<u>Second case: $j = 1$.</u> Then, equating the coefficients in (8.17) we get

$$x = 2\cdot u^2 + v^2 + 4\cdot w^2 + 2\cdot u\cdot w - 4\cdot v\cdot w \ , \tag{8.19}$$

$$\begin{cases} u^2 + 4\cdot v^2 + 18\cdot w^2 - 4\cdot u\cdot v + 8\cdot u\cdot w - 18\cdot v\cdot w = 1 , \\ \qquad 2\cdot v^2 + \ 9\cdot w^2 - 2\cdot u\cdot v + 4\cdot u\cdot w - \ 8\cdot v\cdot w = 0 . \end{cases} \tag{8.20}$$

The first relation of (8.20) can be replaced by

$$u^2 - 2\cdot v\cdot w = 1 . \tag{8.21}$$

Note that u is odd. Put $z = v - 2\cdot w$. Then the second equation of (8.20) yields

$$w^2 = 2\cdot z\cdot (u-z) .$$

First we suppose that z is odd. Then there exist $m, n \in \mathbb{Z}$ such that

$$z = m^2 , \quad u - z = 2\cdot n^2 ,$$

where we use that $u \geqslant 0$ and $(u,w) = 1$. Thus, choosing signs properly,

$$u = m^2 + 2\cdot n^2 , \quad v = m^2 + 4\cdot m\cdot n , \ w = 2\cdot m\cdot n .$$

Substituting this in (8.21) we obtain the Thue equation

$$m^4 - 4\cdot m^3\cdot n - 12\cdot m^2\cdot n^2 + 4\cdot n^4 = 1 . \tag{8.22}$$

In Theorem 8.8(i) below we prove that this equation has only the solutions $\pm(m,n) = (1,0)$, leading to $(u,v,w) = (1,1,0)$, and finally for (8.14) to $(x,\pm y) = (3,4)$.

Secondly we suppose that z is even. Then there exist $m, n \in \mathbb{Z}$ with

$$z = 2\cdot m^2 , \quad u - z = n^2 .$$

Thus, choosing signs properly, we find

$$u = 2\cdot m^2 + n^2 , \quad v = 2\cdot m^2 + 4\cdot m\cdot n , \quad w = 2\cdot m\cdot n .$$

Now, substituting into (8.21), we obtain the Thue equation

$$n^4 - 12\cdot n^2\cdot m^2 - 8\cdot n\cdot m^3 + 4\cdot m^4 = 1 . \tag{8.23}$$

In Theorem 8.8(ii) below we prove that this equation has only the solutions $\pm(m,n) = (0,1), (1,-1), (3,1), (-1,3)$. They lead respectively to $(u,v,w) = (1,0,0), (3,-2,-2), (19,30,6), (11,-10,-6)$, which lead for (8.14) to the solutions $(x,\pm y) = (2,1),\ (10,31),\ (1274,45473),\ (114,1217)$. Thus this result completes the proof of Theorem 8.7, provided the Thue equations (8.22), (8.23) have as their only solutions the pairs (m,n) mentioned above. We now proceed to prove this.

THEOREM 8.8. (i). *The Thue equation*

$$X^4 - 4\cdot X^3\cdot Y - 12\cdot X^2\cdot Y^2 + 4\cdot Y^4 = 1 \tag{8.24}$$

has only the solutions $\pm(X,Y) = (1,0)$.

(ii). *The Thue equation*

$$X^4 - 12\cdot X^2\cdot Y^2 - 8\cdot X\cdot Y^3 + 4\cdot Y^4 = 1 \tag{8.25}$$

has only the solutions $\pm(X,Y) = (1,0),\ (1,-1),\ (1,3),\ (3,-1)$.

Proof. We use the notation and results of Sections 8.2 and 8.3. Let the algebraic numbers ϑ and φ be defined by

$$\vartheta^4 - 12\cdot\vartheta^2 - 8\cdot\vartheta + 4 = 0\ ,\quad \varphi^4 - 4\cdot\varphi^3 - 12\cdot\varphi^2 + 4 = 0\ .$$

Since $\varphi = 2/\vartheta$, it follows that ϑ and φ generate the same field K over $\mathbb{Q}$. In the notation of Section 8.2 we have $n = 4,\ s = 4,\ t = 0$, and $\xi = \vartheta$ or $\xi = \varphi$. Simple computations show that for $\xi = \vartheta,\ \varphi$ we can take

$$Y_0 = 1\ ,\quad C_1 = 0.843\ ,\quad C_2 = 0.589\ ,\quad Y_1 = 2,\quad C_3 = 6.645\ ,$$

$$Y_2^* = 3\ ,\quad \mu_- = \mu_+ = 1\ ,\quad C_4 = 8.3374\ .$$

In these computations we estimated $C_1,\ C_3,\ C_4$ from above and C_2 from below, using the following approximations for the conjugates of ϑ and φ :

$$\begin{aligned}
\vartheta^{(1)} &\cong -1.080\ 286\ 352\ , & \varphi^{(1)} &\cong -1.851\ 360\ 980\ ,\\
\vartheta^{(2)} &\cong 3.722\ 935\ 260\ , & \varphi^{(2)} &\cong 0.537\ 210\ 524\ ,\\
\vartheta^{(3)} &\cong 0.334\ 111\ 716\ , & \varphi^{(3)} &\cong 5.986\ 021\ 747\ ,\\
\vartheta^{(4)} &\cong -2.976\ 760\ 624\ , & \varphi^{(4)} &\cong -0.671\ 871\ 290\ .
\end{aligned}$$

Now we work in the order R of K with $\mathbb{Z}$-basis $\{\ 1,\ \vartheta,\ \frac{1}{2}\cdot\vartheta^2,\ \frac{1}{2}\cdot\vartheta^3\ \}$ (note that $\frac{1}{2}\cdot\vartheta^2$ is an algebraic integer). Note that

$$\varphi = \frac{2}{\vartheta} = 4 + 6\cdot\vartheta - \frac{1}{2}\cdot\vartheta^3 \in R\ .$$

On the other hand, (8.24) and (8.25) are respectively equivalent to

$\mathrm{Norm}_{K/\mathbb{Q}}(X-Y\cdot\vartheta) = 1$ and $\mathrm{Norm}_{K/\mathbb{Q}}(X-Y\cdot\varphi) = 1$, which means that if (X,Y) is a solution of (8.24) or (8.25), then $X - Y\cdot\vartheta$ or $X - Y\cdot\varphi$, respectively, is a unit of the order R. A system of fundamental units of R is given by

$$\varepsilon_1 = 1 + \vartheta\ , \quad \varepsilon_2 = 3 + \vartheta\ , \quad \varepsilon_3 = \tfrac{1}{2}\cdot\vartheta^2\ .$$

We do not prove this fact here. For a proof, see Tzanakis and de Weger [1989[a]], Section III.2 and Appendix I.

Thus the solution of (8.24) and (8.25) is reduced to finding all $(a_1,a_2,a_3) \in \mathbb{Z}^3$ such that the unit $\pm\varepsilon_1^{a_1}\cdot\varepsilon_2^{a_2}\cdot\varepsilon_3^{a_3}$ has the special shape $X - Y\cdot\vartheta$ or $X - Y\cdot\varphi$, respectively. In the notation of Lemma 8.3 we have, after some numerical computations, that we leave to the reader to check, that

$$\min_I N[U_I^{-1}] = 0.634950\ldots\ , \quad \max_I N[U_I^{-1}] = 1.210070\ldots\ ,$$

(here, of course, $I = \{\ 1,\ 2,\ 3,\ 4\ \}$). Therefore we can take in Lemma 8.4

$$C_5 = 1.211\ .$$

Also,

$$C_6 = 6.38771\times10^4\ , \quad Y_2' = 3\ .$$

(The values of C_5 and C_6 are estimated from above.)

Now, relation (8.3) becomes in our case

$$\Lambda = \log\left|\frac{\xi^{(i_0)}-\xi^{(j)}}{\xi^{(i_0)}-\xi^{(k)}}\right| + \sum_{i=1}^{3} a_i\cdot\log\left|\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right|\ , \tag{8.26}$$

where $\xi = \vartheta$ or φ. As mentioned in Section 8.2, once i_0 is fixed, we can choose $j,\ k$ arbitrarily. Thus we can choose

$$\begin{cases} j = 3,\ k = 4 \quad \text{if} \quad i_0 = 1 \quad \text{or} \quad 2\ , \\ j = 1,\ k = 2 \quad \text{if} \quad i_0 = 3 \quad \text{or} \quad 4\ . \end{cases} \tag{8.27}$$

Therefore, for each $\xi \in \{\ \vartheta,\ \varphi\ \}$ we have four possibilities for Λ. For each of these eight cases we have, as will be shown below,

$$C_7 = 5.71\times10^{38}\ , \quad C_8 = 6.17\ ,$$

and therefore, by Lemma 8.5, if $|Y| > 3$, then for $A = \max_{1\leqslant i\leqslant 3} |a_i|$ we have the upper bound $C_9 = 3.26\times10^{40}$. As is easily checked, the only solutions of either (8.24) or (8.25) with $|Y| \leqslant 3$ are those listed in the statement of the theorem. Therefore we may assume that $|Y| > 3$, so that

$$A < 3.26\times10^{40} .$$

Before we apply the reduction method of Section 3.8 we show that the application of Lemma 2.4 yields the above constants C_7, C_8 . We apply this result in the case of Λ given by (8.26). In this case, we compute the V_i's for the various α_i's appearing in Λ , as follows. If $\alpha_i = |\varepsilon_i^{(k)}/\varepsilon_i^{(j)}|$ for $i = 1, 2, 3$, then α_i is a unit and hence a_0 (appearing in the computation of $h(\alpha_i)$) is equal to 1. Clearly, every conjugate of α_i is in absolute value less than

$$H_i = \frac{\max\limits_{1\leqslant h\leqslant 4}|\varepsilon_i^{(h)}|}{\min\limits_{1\leqslant h\leqslant 4}|\varepsilon_i^{(h)}|} ,$$

and $H_i \geqslant 1$. Therefore, $h(\alpha_i) \leqslant H_i$, and we can take

$$V_i = \max \Big(\log H_i, \ |\log|\varepsilon_i^{(k)}/\varepsilon_i^{(j)}|| \Big) .$$

Since the latter term equals the logarithm of either $|\varepsilon_i^{(k)}/\varepsilon_i^{(j)}|$ or its inverse, it follows that

$$V_i = \log H_i .$$

If $\alpha_i = |\xi^{(i_0)}-\xi^{(j)}|/|\xi^{(i_0)}-\xi^{(k)}|$, then all conjugates of α_i are in absolute value less than C_3 . Therefore, $h(\alpha_i) \leqslant (\log a_0)/d + \log C_3$, where a_0 and d are as in the definition of $h(\alpha)$ for $\alpha = \alpha_i$. An upper bound for a_0 can be computed as follows. Consider the algebraic numbers $\chi_{ih} = \frac{1}{2}\cdot(\xi^{(i)}-\xi^{(h)})$ for $i, h \in \{ 1, \ldots, 4 \}$ with $i \neq h$. It can be checked that the numbers χ_{ih} are algebraic integers for $\xi = \vartheta$ or φ . Now, for each permutation $\sigma = (\sigma_1\sigma_2\sigma_3\sigma_4) \in S_4$ we consider the number $\chi(\sigma) = \chi_{\sigma_1\sigma_2}/\chi_{\sigma_1\sigma_3}$ (independent of σ_4), and the polynomial

$$P(X) = \prod_{\sigma\in S_4} \big(X - \chi(\sigma) \big) .$$

Consider also the number

$$\Delta = \prod_{1\leqslant i<h\leqslant 4} \chi_{ih} .$$

Note that

$$\Delta^2 = \frac{1}{2^{12}} \cdot \prod_{1 \leqslant i < h \leqslant 4} (\xi_i - \xi_h)^2 = \frac{1}{2^{12}} \cdot D ,$$

where D is the discriminant of the defining polynomial of ξ , and therefore $\Delta^2 = 229$. On the other hand, the coefficients of $P(X)$ are up to the sign the elementary symmetric functions of $\chi(\sigma)$ for $\sigma \in S_4$, and so they are symmetrical expressions of the $\xi^{(i)}$'s with rational coefficients. This means that $P(X) \in \mathbb{Q}[X]$. On the other hand, by the definition of Δ , any coefficient of $P(X)$ multiplied by Δ^4 is a polynomial of the χ_{ih}'s with coefficients in $\mathbb{Z}$ and therefore it is an algebraic integer. Combine this with the fact that $P(X) \in \mathbb{Q}[X]$ to see that $229^2 \cdot P(X) \in \mathbb{Z}[X]$. Hence, since α_i is a root of $P(X)$, its leading coefficient a_0 is at most 229^2 . To conclude, we have $h(\alpha_i) \leqslant 2 \cdot (\log 229)/d + \log C_3$ and it is clear that $|\log \alpha_i|/d \leqslant \log C_3$. Since $\alpha_i \notin \mathbb{Q}$ we have $d \geqslant 2$, so we can take

$$V_i = \log 229 + \log C_3 .$$

Simple computations now show that

$$\log H_1 = 4.074586\ldots , \quad \log H_2 = 5.667432\ldots ,$$

$$\log H_3 = 4.821584\ldots ,$$

$$\log C_3 = 1.262065\ldots \quad \text{if} \quad \xi = \vartheta ,$$

$$\log C_3 = 1.893823\ldots \quad \text{if} \quad \xi = \varphi ,$$

$$\log 229 + \log C_3 \leqslant 7.327545\ldots .$$

Therefore we apply Lemma 2.4 (Waldschmidt) with $n = 4,\ D \leqslant 24,\ e(n) = 73,$

$$\alpha_1 = \left|\frac{\varepsilon_1^{(k)}}{\varepsilon_1^{(j)}}\right| , \quad \alpha_2 = \left|\frac{\varepsilon_3^{(k)}}{\varepsilon_3^{(j)}}\right| , \quad \alpha_3 = \left|\frac{\varepsilon_2^{(k)}}{\varepsilon_2^{(j)}}\right| , \quad \alpha_4 = \left|\frac{\xi^{(i_0)} - \xi^{(j)}}{\xi^{(i_0)} - \xi^{(k)}}\right| ,$$

for $\xi = \vartheta$ or φ , and $b_1 = a_1$, $b_2 = a_3$, $b_3 = a_2$, $b_4 = 1$, $B = A$, $V_1 = \log H_1$, $V_2 = \log H_3$, $V_3 = V_3^+ = \log H_2$, $V_4 = V_4^+ = \log 229 + \log C_3$. Thus we find that

$$|\Lambda| > \exp\bigl(-C_7 \cdot (\log A + C_8)\bigr) ,$$

with $C_7 = 5.71 \times 10^{38}$ and $C_8 = 6.17$.

We have now to apply the reduction process described in Section 3.7. In our situation we have to solve (8.9) with

$$K_1 = C_6 = 6.38771\times 10^4 \ , \quad K_2 = \frac{n}{C_5} = \frac{4}{1.211} > 3.303 \ , \quad K_3 = 3.26\times 10^{40}$$

(K_2 is estimated from below), and

$$\Lambda = \delta + a_1\cdot\mu_1 + a_2\cdot\mu_2 + a_3\cdot\mu_3 \ ,$$

where for δ and the μ_i's we have, in view of (8.26) and (8.27):

$$\begin{cases} \delta = \delta_1 := \log\left|\dfrac{\xi^{(1)}-\xi^{(3)}}{\xi^{(1)}-\xi^{(4)}}\right| \quad \text{or} \quad \delta = \delta_2 := \log\left|\dfrac{\xi^{(2)}-\xi^{(3)}}{\xi^{(2)}-\xi^{(4)}}\right| \ , \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{where} \quad \xi = \vartheta \quad \text{or} \quad \varphi \ , \\ \mu_i = \log\left|\dfrac{\varepsilon_i^{(4)}}{\varepsilon_i^{(3)}}\right| \ , \quad \text{for} \quad i = 1,\ 2,\ 3 \ , \end{cases} \tag{8.28}$$

or

$$\begin{cases} \delta = \delta_3 := \log\left|\dfrac{\xi^{(3)}-\xi^{(1)}}{\xi^{(3)}-\xi^{(2)}}\right| \quad \text{or} \quad \delta = \delta_4 := \log\left|\dfrac{\xi^{(4)}-\xi^{(1)}}{\xi^{(4)}-\xi^{(2)}}\right| \ , \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{where} \quad \xi = \vartheta \quad \text{or} \quad \varphi \ , \\ \mu_i = \log\left|\dfrac{\varepsilon_i^{(2)}}{\varepsilon_i^{(1)}}\right| \ , \quad \text{for} \quad i = 1,\ 2,\ 3 \ . \end{cases} \tag{8.29}$$

Numerical details are given in the preprint version of Tzanakis and de Weger [1989[a]] (to be obtained from the author). We take $c_0 = 10^{140}$, and we work with the lattice with associated matrix

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ [c_0\cdot\mu_1] & [c_0\cdot\mu_2] & [c_0\cdot\mu_3] \end{pmatrix} .$$

Note that in each of the four cases of (8.28) (resp. (8.29)) we have the same lattice, Γ_1 (resp. Γ_2), say. In each case $\delta \neq 0$, and we had no numerical evidence that the μ_i's are $\mathbb{Q}$-dependent. Therefore we worked as in case (ii) of Section 8.4.

For each Γ_i we have applied the integral version of the L^3-algorithm, and each time we have computed the integral 3×3-matrices $\mathcal{B}$, $\mathcal{U}$, $\mathcal{U}^{-1}$, as defined in Section 3.7. In our cases, the coordinates of the vectors of the reduced bases (i.e. the elements of $\mathcal{B}$) turned out to have 46 to 48 digits, i.e. the lengths of the reduced basis vectors are of the size of $c_0^{1/3}$, as expected.

In each of the eight cases we computed the coordinates s_1, s_2, s_3 of

$$\underline{x} = \begin{pmatrix} 0 \\ 0 \\ -[c_0\cdot\delta] \end{pmatrix}$$

with respect to the reduced basis $\underline{b}_1$, $\underline{b}_2$, $\underline{b}_3$ of the lattice. From our computations we found

$$|\underline{b}_1| > 3.247\times10^{46} \quad \text{in the case of lattice } \Gamma_1 ,$$

$$|\underline{b}_1| > 4.846\times10^{46} \quad \text{in the case of lattice } \Gamma_2 ,$$

$$\|s_3\| > 0.029 \quad \text{in all 8 cases.}$$

This means that in view of Lemma 3.5, in all cases $i_0 = 3$, and

$$\ell(\Gamma_i,\underline{x}) > 0.029\cdot\tfrac{1}{2}\cdot3.247\times10^{46} > 4.708\times10^{44} .$$

Then the assumptions of Lemma 3.10 are fulfilled with $n = 3$, $\gamma = 1$, $C = c_0$, $c = K_1$, $\delta = K_2$, $X_0 = X_1 = K_3$, since $\sqrt{27}\cdot K_3 < 1.112\times10^{40}$, which implies

$$A < \frac{1}{3.303}\cdot\log\left(10^{140}\cdot6.38771\times10^{4}/3.26\times10^{40}\right) < 72.8 .$$

It follows that $A \leqslant 72$. We repeat the procedure with $K_3 = 72$ and $c_0 = 10^{12}$. We found from our computations

$$|\underline{b}_1| > 1.293\times10^{4} \quad \text{in the case of lattice } \Gamma_1 ,$$

$$|\underline{b}_1| > 1.092\times10^{4} \quad \text{in the case of lattice } \Gamma_2 ,$$

$$\|s_3\| > 0.143 \quad \text{in all 8 cases.}$$

This means that in view of Lemma 3.5, in all cases $i_0 = 3$, and

$$\ell(\Gamma_i,\underline{x}) > 0.143\cdot\tfrac{1}{2}\cdot1.092\times10^{4} > 7.807\times10^{2} .$$

Then the assumptions of Lemma 3.10 are fulfilled, since $\sqrt{27}\cdot K_3 < 3.742\times10^{2}$, which implies

$$A < \frac{1}{3.303}\cdot\log\left(10^{12}\cdot6.38771\times10^{4}/72\right) < 10.5 .$$

It follows that $A \leqslant 10$. We enumerated all remaining possibilities, and found no other solutions of (8.24) and (8.25) than those mentioned. □

The computations for the proof of Theorem 8.8 took 35 sec.

8.6. The Thue-Mahler equation, an outline.

Let $F(X,Y)$ be as in Section 8.1. Let $p_1, \ldots, p_s$ be fixed distinct prime numbers. The diophantine equation

$$F(X,Y) = \pm \prod_{i=1}^{s} p_i^{n_i}$$

in the variables $X, Y \in \mathbb{Z}$, $n_1, \ldots, n_s \in \mathbb{N}_0$, with $(X,Y) = 1$, is known as a Thue-Mahler equation. It was proved by Mahler [1933] that this equation has only finitely many solutions, and by Coates [1970] that they can, at least in principle, be determined effectively, since an effectively computable upper bound for the variables can be derived from the p-adic theory of linear forms in logarithms. For the history of this equation we refer to Shorey and Tijdeman [1986], Chapter 7.

We believe that it is possible to solve Thue-Mahler equations, not only in principle, but in practice. This can be done by reducing the above mentioned upper bounds, using a combination of real and p-adic computational diophantine approximation techniques, based on the L^3-algorithm for reducing bases of lattices (cf. Sections 3.7 and 3.8 for the real case, 3.11 and 3.12 for the p-adic case, Section 1.5 for a short outline of how to combine the real and p-adic techniques, and Sections 4.8 and 6.4 for some explicit examples of such combined techniques). The method can be considered as a p-adic analogue of the method for solving Thue equations, on which we reported in the preceding sections.

Such an idea (but without using the L^3-algorithm) was used by Agrawal, Coates, Hunt and van der Poorten [1980], who solved the equation

$$X^3 - X^2 \cdot Y + X \cdot Y^2 + Y^3 = \pm 11^n .$$

This is to the author's knowledge the only example in the literature where a Thue-Mahler equation has been solved by the Gelfond-Baker method. Other methods may apply as well for solving Thue-Mahler equations. For example,

$$X^3 + 3 \cdot Y^3 = 2^n ,$$

has been solved by Tzanakis [1984] by a different method. The advantage of the Gelfond-Baker method above many other ideas is that it works in principle for any Thue-Mahler equation, because it is not very much dependent on the parameters of the particular equation that one wants to solve.

Both examples of Thue-Mahler equations mentioned above are of the simplest kind, in view of the fact that the cubic field $\mathbb{Q}(\vartheta)$, where ϑ is a root of $F(x,1) = 0$, has only one fundamental unit, and there occurs only one prime. Therefore it is sufficient to use two-dimensional real continued fractions and one-dimensional p-adic continued fractions, instead of the more complicated L^3-algorithm (which anyway was not yet available in 1980, when Agrawal, Coates, Hunt and van der Poorten did their work). With the use of the L^3-algorithm the method can in principle be extended to the general situation, where there are more than one fundamental units, and more than one primes. In a forthcoming publication, Tzanakis and the present author plan to give details and worked-out examples (Tzanakis and de Weger [1989[b]]).

References.

After each reference we mention in brackets the section(s) in which the reference occurs.

Agrawal, M.K., Coates, J.H., Hunt, D.C. and van der Poorten, A.J. [1980], Elliptic curves of conductor 11, *Math. Comp.* **35**, 991-1002. (3.8;3.10; 8.6)

Alex, L.J. [1976], Diophantine equations related to finite groups, *Comm. Algebra* **4**, 77-100. (6.1;6.5)

Alex, L.J. [1985[a]], On the diophantine equation $1 + 2^a = 3^b 5^c + 2^d 3^e 5^f$, *Math. Comp.* **44**, 267-278. (1.1;5.4)

Alex, L.J. [1985[b]], On the diophantine equation $1 + 2^a = 3^b 7^c + 2^d 3^e 7^f$, *Arch. Math.* **45**, 538-545. (1.1;5.4)

Babai, L. [1986], On Lovász lattice reduction and the nearest lattice point problem, *Combinatorica* **6**, 1-13. (1.4;3.4)

Bachman, G. [1964], *Introduction to p-adic Numbers and Valuation Theory*, Academic Press, New York and London. (2.3)

Baker, A. [1966], Linear forms in the logarithms of algebraic numbers, *Mathematika* **13**, 204-216. (2.4)

Baker, A. [1968], Contributions to the theory of diophantine equations, I, On the representation of integers by binary forms, II, The diophantine equation $y^2 = x^3 + k$, *Phil. Trans. R. Soc. London*, A **263**, 173-208. (8.1)

Baker, A. [1972], A sharpening of the bounds for linear forms in logarithms I, *Acta Arith.* **21**, 117-129. (1.2)

Baker, A. [1977], The theory of linear forms in logarithms, *Transcendence Theory: Advances and Applications*, A. Baker (ed.), Academic Press, London, pp. 1-27. (2.4)

Baker, A. and Davenport, H. [1969], The equations $3x^2 - 2 = y^2$ and $8x^2 - 7 = z^2$, *Q. Jl. Math. Oxford* (2) **20**, 129-137. (1.1;1.4;3.3)

Berwick, W.E.H. [1932], Algebraic number fields with two independent units, *Proc. London Math. Soc.* **34**, 360-378. (8.2)

Beukers, F. [1981], On the generalized Ramanujan-Nagell equation, *Acta Arith.* I: **38**, 389-410; II: **39**, 113-123. (1.1;4.1)

Billevič, K.K. [1956], On the units of algebraic fields of third and fourth degree (Russian), *Mat. Sbornik* Vol. **40**, **82**, 123-137. (8.2)

Billevič, K.K. [1964], A theorem on the units of algebraic fields of n-th degree (Russian), *Mat. Sbornik* Vol. **64**, **106**, 145-152. (8.2)

Blass, J., Glass, A.M.W., Meronk, D.B. and Steiner, R.P. [1987[a]], Practical solutions to Thue equations of degree 4 over the rational integers, *Preprint*, Bowling Green State University. (3.8;8.1)

Blass, J., Glass, A.M.W., Meronk, D.B. and Steiner, R.P. [1987[b]], Practical solutions to Thue equations over the rational integers, *Preprint*, Bowling Green State University. (3.8;8.1;8.4)

Blass, J., Glass, A.M.W., Manski, D.K., Meronk, D.B. and Steiner, R.P. [1988[a]], Constants for lower bounds for linear forms in the logarithms of algebraic numbers I: The general case, *Preprint*, Bowling Green State University. (2.4)

Blass, J., Glass, A.M.W., Manski, D.K., Meronk, D.B. and Steiner, R.P. [1988[b]], Constants for lower bounds for linear forms in the logarithms of algebraic numbers II: The homogeneous rational case, *Preprint*, Bowling Green State University. (2.4)

Borevich, Z.I. and Shafarevich, I.R. [1966], *Number Theory*, Academic Press, New York. (2.1;7.3)

Bremner, A., Calderbank, R., Hanlon, P., Morton, P. and Wolfskill, J. [1983], Two-weight ternary codes and the equation $y^2 = 4\times 3^\alpha + 13$, *J. Number Theory* **16**, 212-234. (4.1)

Brenner, J.L. and Foster, L.L. [1982], Exponential diophantine equations, *Pacific J. Math.* **101**, 263-301. (6.1)

Brent, R.P. [1978], A Fortran multiple-precision arithmetic package, *ACM Trans. Math. Software* **4**, 57-70; and: Algorithm 524. MP, a Fortran multiple-precision arithmetic package, *ACM Trans. Math. Software* **4**, 71-81. (2.5)

Brentjes, A.J. [1981], *Multi-dimensional Continued Fraction Algorithms*, MC Tract 145, Centr. Math. Comput. Sci., Amsterdam. (1.3;3.4)

Brown, E. [1985], Sets in which $xy + k$ is always a square, *Math. Comp.* **45**, 613-620. (1.1)

Buchmann, J. [1985], The generalized Voronoi-algorithm in totally real algebraic number fields, *Proceedings EUROCAL '85*, Linz, Austria, Vol. 2, Lecture Notes in Comput. Sci. 204, Springer Verlag, Berlin, pp. 479-486. (8.2)

Buchmann, J. [1986], A generalization of Voronoi's unit algorithm I & II, *J. Number Theory* **20**, 177-191 & 192-209. (8.2;8.5)

Cassels, J.W.S. [1957], *An Introduction to Diophantine Approximation*, Cambridge University Press, Cambridge. (1.3)

Cherubini, J.M. and Walliser, R.V. [1987], On the computation of all imaginary quadratic fields of class number one, *Math. Comp.* **49**, 295-300. (3.2)

Coates, J. [1969], An effective p-adic analogue of a theorem of Thue, *Acta Arith.* **15**, 279-305. (2.4;6.1)

Coates, J. [1970], An effective p-adic analogue of a theorem of Thue II: The greatest prime factor of a binary form, *Acta Arith.* **16**, 399-412. (2.4; 6.1;8.6)

Delone, B.N. and Faddeev, D.K. [1964], *The theory of irrationalities of the third degree*, Transl. of Math. Monogr., Vol 10, A.M.S., Providence R.I. (8.5)

Ellison, W.J. [1971[a]], Recipes for solving diophantine problems by Baker's method, Sém. Théorie des Nombres, Université de Bordeaux I, 1970-1, Lab. Th. Nombr. C.N.R.S., Exp. 11, 10 pp. (1.4;3.8)

Ellison, W.J. [1971[b]], On a theorem of S. Sivasankaranarayana Pillai, Sém. Théorie des Nombres, Université de Bordeaux I, 1970-1, Lab. Th. Nombr. C.N.R.S., Exp. 12, 10 pp. (3.2;5.1)

Ellison, W.J., Ellison, F., Pesek, J., Stahl, C.E. and Stall, D.S. [1972], The diophantine equation $y^2 + k = x^3$, *J. Number Theory* **4**, 107-117. (3.3;8.1)

Evertse, J.-H. [1983], *Upper Bounds for the Numbers of Solutions of Diophantine Equations*, MC Tract 168, Centr. Math. Comput. Sci., Amsterdam. (1.1)

Evertse, J.-H., Győry, K., Stewart, C.L. and Tijdeman, R. [1988], S-unit equations and their applications, *New advances in transcendence theory (Proc. Symp. Durham July 1986)*, A. Baker (ed.), Cambridge University Press, Cambridge, pp. 110-174. (1.1)

Faltings, G. [1983], Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, *Invent. Math.* **73**, 349-366. (1.1)

Fincke, U. and Pohst, M. [1985], Improved methods for calculating vectors of short length in a lattice, including a complexity analysis, *Math. Comp.* **44**, 463-471. (3.6)

Gaál, I. [1988], On the resolution of inhomogeneous norm form equations in two dominating variables, *Math. Comp.* **51**, 359-373. (3.3)

Grinstead, C.M. [1978], On a method of solving a class of diophantine equations, *Math. Comp.* **32**, 936-940. (1.1)

Grosswald, E. [1984], *Topics from the Theory of Numbers*, 2nd. ed., Birkhäuser, Boston. (1.1)

Hardy, G.H. and Wright, E.M. [1979], *An Introduction to the Theory of Numbers*, (5th ed.), Oxford University Press, Oxford. (1.3;3.2)

Hasse, H. [1966], Über eine diophantische Gleichung von Ramanujan-Nagell und ihre Verallgemeinerung, *Nagoya Math. J.* **27**, 77-102. (4.1)

Hunt, D.C. and van der Poorten, A.J., Solving diophantine equations $x^2 + d = a^u$, *unpublished.* (3.2;4.1)

Kiss, P. [1979], Zero terms in second order linear recurrences, *Math. Sem. Notes Kobe Univ.* (Japan) **7**, 145-152. (4.3)

Knuth, D.E. [1981], *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms*, (2nd ed.), Addison-Wesley, Reading Mass. (2.5)

Koblitz, N. [1977], *p-adic Numbers, p-adic Analysis, and Zeta-functions*, Springer Verlag, New York. (2.3)

Koblitz, N. [1980], *p-adic Analysis: a Short Course on Recent Work*, Cambridge University Press, Cambridge. (2.3)

Koksma, J.F. [1937], Diophantische Approximationen, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Vol. **4**, Springer Verlag, pp. 407-571. (1.3)

Lagarias, J.C. and Odlyzko, A.M. [1985], Solving low-density subset sum problems, *J. Assoc. Comput. Mach.* **32**, 229-246. (3.4)

Langevin, M. [1976], Quelques applications de nouveaux résultats de van der Poorten, Sém. Delange-Pisot-Poitou 1975/76, Paris, Exp. G12, 11 pp. (1.1)

Lehmer, D.H. [1964], On a problem of Störmer, *Illinois J. Math.* **8**, 57-79. (4.9;5.1)

Lenstra, A.K. [1984], *Polynomial-time Algorithms for the Factorization of Polynomials*, Dissertation, University of Amsterdam. (1.4;3.4;3.5)

$\mathcal{LLL}$ = Lenstra, A.K., Lenstra, H.W. Jr. and Lovász, L. [1982], Factoring polynomials with rational coefficients, *Math. Ann.* **261**, 515-534. (1.4;3.4;3.5)

Loxton, J.H., Mignotte, M., van der Poorten, A.J. and Waldschmidt, M. [1987], A lower bound for linear forms in the logarithms of algebraic numbers, *C.R. Math. rep. Acad. Sci. Canada* **11**, 119-124. (2.4)

Lutz, É. [1951], *Sur les approximations diophantiennes linéaires P-adiques*, Thèse, Université de Strasbourg. (1.3)

MacWilliams, F.J. and Sloane, N.J.A. [1977], *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam. (4.1)

Mahler, K. [1933], Zur Approximation algebraischer Zahlen I: Über den grössten Primteiler binärer Formen, *Math. Ann.* **107**, 691-730. (6.1;8.6)

Mahler, K. [1934], Eine arithmetische Eigenschaft der rekurrierenden Reihen, *Mathematika* B (Leiden) **3**, 153-156. (4.1)

Mahler, K. [1935], Über den grössten Primteiler spezieller Polynome zweiten Grades, *Arch. Math. Naturvid.* B **41**, 3-26. (4.9;5.1)

Mahler, K. [1961], *Lectures on Diophantine Approximations I: g-adic Numbers and Roth's Theorem*, University of Notre Dame Press, Notre Dame. (3.10)

Masser, D.W. [1985], Open Problems, *Proc. Symp. Analytic Number Th.*, W.W.L. Chen (ed.), London, Imperial College. (6.6)

Mignotte, M. [1984[a]], On the automatic resolution of certain diophantine equations, *Proceedings of EUROSAM 84*, Lecture Notes in Comput. Sci. 174, Springer Verlag, Berlin, pp. 378-385. (4.1)

Mignotte, M. [1984[b]], Une nouvelle résolution de l'équation $x^2 + 7 = 2^n$, *Rend. Sem. Fac. Sci. Univ. Cagliari* **54**, Fasc. 2, 41-43. (4.1)

Mignotte, M. [1985], $P(x^2+1) \geqslant 17$ si $x \geqslant 240$, *C. R. Acad. Sci. Paris* **301**, Série I, No. 13, 661-664. (4.9)

Mignotte, M. and Waldschmidt, M. [1978], Linear forms in two logarithms and Schneider's method, *Math. Ann* **231**, 241-267. (2.4)

Mignotte, M. and Waldschmidt, M. [1988], Linear forms in two logarithms and Schneider's method (II), *Preprint*, I.R.M.A., Université Louis Pasteur, Strasbourg. (2.4)

Mohanty, S.P. [1988], Integer points of $y^2 = x^3 - 4x + 1$, *J. Number Theory* **30**, 86-93. (8.5)

Nagell, T. [1948], Løsning Oppg. 2, 1943, s.29 (Norwegian), *Norsk Mat. Tidsskr.* **30**, 62-64. (1.1;4.1;4.9)

Odlyzko, A.M. and te Riele, H.J.J. [1985], Disproof of the Mertens conjecture, *J. reine angew. Math.* **357**, 138-160. (3.4)

Pethö, A. [1983], Full cubes in the Fibonacci sequence, *Publ. Math. Debrecen* **30**, 117-127. (1.1;8.1).

Pethö, A. [1985], On the solution of the diophantine equation $G_n = p^z$, *Proceedings EUROCAL '85*, Linz, Austria, Vol 2, Lecture Notes in Comput. Sci. 204, Springer Verlag, Berlin, pp. 503-512. (4.1;4.7)

Pethö, A. and Schulenberg, R. [1987], Effektives Lösen von Thue Gleichungen, *Publ. Math. Debrecen* **34**, 189-196. (3.8;8.1)

Pethö, A. and de Weger, B.M.M. [1986], Products of prime powers in binary recurrence sequences I: The hyperbolic case, with an application to the generalized Ramanujan-Nagell equation, *Math. Comp.* **47**, 713-727. (1.1; 2.2;4)

Philippon, P. and Waldschmidt, M. [1988], Lower bounds for linear forms in logarithms, *New advances in transcendence theory (Proc. Symp. Durham July 1986)*, A. Baker (ed.), Cambridge University Press, Cambridge, pp. 280-312. (2.4)

Pinch, R.G.E. [1984], Elliptic curves with good reduction away from 2, *Math. Proc. Camb. Phil. Soc.* **96**, 25-38. (7.1).

Pinch, R.G.E. [1988], Simultaneous Pellian equations, *Math. Proc. Camb. Phil. Soc.* **103**, 35-46. (1.1)

Pohst, M. and Zassenhaus, H. [1982], On effective computation of fundamental units I & II, *Math. Comp.* **38**, 275-291 & 293-329. (8.2)

van der Poorten, A.J. [1977], Linear forms in logarithms in the p-adic case, *Transcendence Theory: Advances and Applications*, A. Baker (ed.), Academic Press, London, pp. 29-57. (1.2;2.4;6.2)

Rumsey, H. Jr. and Posner, E.C. [1964], On a class of exponential equations, *Proc. Am. Math. Soc.* **15**, 974-978. (4.9;6.1)

Schinzel, A. [1967], On two theorems of Gelfond and some of their applications, *Acta Arith.* **13**, 177-236. (2.4;4.1)

Schmidt, W.M. [1988], The number of solutions of Thue equations, *New advances in transcendence theory (Proc. Symp. Durham July 1986)*, A. Baker (ed.), Cambridge University Press, Cambridge, pp. 337-346. (1.1)

Setzer, B. [1975], Elliptic curves of prime conductor, *J. London Math. Soc.* **10**, 367-378. (7.1)

Shorey, T.N., van der Poorten, A.J., Tijdeman, R. and Schinzel, A. [1977], Applications of the Gel'fond-Baker method to diophantine equations, *Transcendence Theory: Advances and Applications*, A. Baker (ed.), Academic Press, London, pp. 59-77. (1.1)

Shorey, T.N. and Tijdeman, R. [1986], *Exponential Diophantine Equations*, Cambridge University Press, Cambridge. (1.1;1.2;4.2;5.1;6.1;8.1;8.6)

Sprindžuk, V.G. [1969], Effective estimates in 'ternary' exponential diophantine equations (Russian), *Dokl. Akad. Nauk BSSR* **12**, 293-297. (6.1)

Steiner, R.P. [1977], A theorem on the Syracuse problem, *Proc. Seventh Manitoba Conf. Numer. Math. Comp.*, pp. 553-559. (3.2)

Steiner, R.P. [1986], On Mordell's equation $y^2 - k = x^3$: a problem of Stolarsky, *Math Comp.* **46**, 703-714. (3.3;8.1)

Stewart, C.L. and Tijdeman, R. [1986], On the Oesterlé-Masser conjecture, *Monatsh. Math.* **102**, 251-257. (6.6)

Størmer, C. [1897], Quelques théorèmes sur l'équation de Pell $x^2 - Dy^2 = \pm 1$ et leurs applications, *Vid. Skr. I Math. Natur. Kl.* (Christiana), 1897, No. 2, 48 pp. (4.9;5.1)

Stroeker, R.J. and Tijdeman, R. [1982], Diophantine equations, (with an Appendix by P.L. Cijsouw, A. Korlaar and R. Tijdeman), *Computational Methods in Number Theory*, H.W. Lenstra and R. Tijdeman (eds.), MC Tract 155, Centr. Math. Comp. Sci., Amsterdam, pp. 321-369. (1.1;3.2;5.1;5.4)

Thue, A. [1909], Über Annäherungswerten algebraischer Zahlen, *J. reine angew. Math.* **135**, 284-305. (8.1)

Tijdeman, R. [1973], On integers with many small prime factors, *Compositio Math.* **26**, 319-330. (5.1)

Tijdeman, R. [1976], On the equation of Catalan, *Acta Arith.* **29**, 197-209. (1.1)

Tijdeman, R. [1985], On the Fermat-Catalan equation, *Jahresber. Deutsche Math. Verein.* **87**, 1-18. (1.1)

Tijdeman, R. [1989], Diophantine equations and diophantine approximations, *Proceedings NATO Advanced Study Institute on Number Theory and Applications, Banff, April-May 1988*, R.A. Mollin (ed.), NATO ASI Series, Reidel, Dordrecht, to appear. (6.6)

Tijdeman, R. and Wang, L. [1988], Sums of products of powers of given prime numbers, *Pacific J. Math.* **132**, 177-193. (1.1;5.4)

Tzanakis, N. [1983], On the diophantine equation $y^2 - D = 2^k$, *J. Number Theory* **17**, 144-164. (4.1)

Tzanakis, N. [1984], The complete solution in integers of $X^3 + 2Y^3 = 2^n$, *J. Number Theory* **19**, 203-208. (8.6)

Tzanakis, N. [1989], On the practical solution of the Thue equation, an outline, *Proceedings Colloquium on Number Theory, Budapest, July 1987*, K. Györy (ed.), North Holland, Amsterdam, to appear. (8.1)

Tzanakis, N. and de Weger, B.M.M. [1989[a]], On the practical solution of the Thue equation, *J. Number Theory* **31**, 99-132. Preprint version with numerical details: *Memorandum* No. **668**, Faculty of Applied Mathematics, University of Twente, October 1987. (1.1;8)

Tzanakis, N. and de Weger, B.M.M. [1989[b]], Solving the diophantine equation $x^3 - 3\cdot x\cdot y^2 - y^3 = \pm 3^{n_0}\cdot 17^{n_1}\cdot 19^{n_2}$, to appear. (8.6)

Tzanakis, N. and Wolfskill, J. [1986], On the diophantine equation $y^2 = 4q^n + 4q + 1$, *J. Number Theory* **23**, 219-237. (4.1)

Tzanakis, N. and Wolfskill, J. [1987], The diophantine equation $x^2 = 4q^{a/2} + 4q + 1$ with an application in coding theory, *J. Number Theory* **26**, 96-116. (4.1)

Vojta, P. [1987], *Diophantine Approximations and Value Distribution Theory*, Lecture Notes in Mathematics 1239, Springer Verlag, Berlin. (6.6)

Wagstaff, S.S. Jr. [1979], Solution of Nathanson's exponential congruence, *Math. Comp.* **33**, 1097-1100. (3.9)

Wagstaff, S.S. Jr. [1981], The computational complexity of solving exponential congruences, *Congressus Numerantium*, University of Winnipeg, Canada, Vol. **31**, pp. 275-286. (3.9)

Waldschmidt, M. [1980], A lower bound for linear forms in logarithms, *Acta Arith.* **37**, 257-283. (2.4)

de Weger, B.M.M. [1986^a], Approximation lattices of p-adic numbers, *J. Number Theory* **24**, 70-88. (1.3;1.4;3.10)

de Weger, B.M.M. [1986^b], Products of prime powers in binary recurrence sequences II: The elliptic case, with an application to a mixed quadratic-exponential equation, *Math. Comp.* **47**, 729-739. (1.1;4)

de Weger, B.M.M. [1987], Solving exponential diophantine equations using lattice basis reduction algorithms, *J. Number Theory* **26**, 325-367. Erratum: *J. Number Theory* **31** [1989], 88-89. (1.1;5;6)

de Weger, B.M.M. [1989^a], On the practical solution of Thue-Mahler equations, an outline, *Proceedings Colloquium on Number Theory, Budapest 1987*, K. Győry (ed.), North Holland, Amsterdam, to appear. (1.1;8.6)

de Weger, B.M.M. [1989^b], A diophantine equation of Antoniadis, *Proceedings NATO Advanced Study Institute on Number Theory and Applications, Banff, April-May 1988*, R.A. Mollin (ed.), Reidel, Dordrecht, to appear. (3.3;8.1)

Wüstholz, G. [1988], A new approach to Baker's theorem on linear forms in logarithms III, *New advances in transcendence theory (Proc. Symp. Durham July 1986)*, A. Baker (ed.), Cambridge University Press, Cambridge, pp. 399-410. (2.4)

Yu, K.R. [1987], Linear forms in the p-adic logarithms, *Report* **MPI/87-20**, Max Planck Institut für Mathematik, Bonn. To appear in *Acta Arith.* (1.2;2.4;6.2)

Yu, K.R. [1988], Linear forms in logarithms in the p-adic case, *New advances in transcendence theory (Proc. Symp. Durham July 1986)*, A. Baker (ed.), Cambridge University Press, Cambridge, pp. 411-434. (2.4)

Yu, K.R. [1989], Linear forms in p-adic logarithms, to appear. (2.4;7.6)