Johannes Blömer

Simplifying Expressions Involving Radicals

Dissertation
Simplifying Expressions Involving Radicals

Dissertation
zur Erlangung des Doktorgrades

vorgelegt am Fachbereich Mathematik
der Freien Universität Berlin
1993

von
Johannes Blömer

verteidigt am 29. Januar 1993

Betreuer  Prof. Dr. Helmut Alt
          Freie Universität Berlin

Zweitgutachter Prof. Dr. Chee Yap
              Courant Institute, New York University
## Contents

1 Introduction  
2 Basics from Algebra and Algebraic Number Theory  
3 The Structure of Radical Extensions  
4 Radicals over the Rational Numbers  
   4.1 Linear Dependence of Radicals over the Rational Numbers  
   4.2 Comparing Sums of Square Roots  
5 Linear Dependence of Radicals over Algebraic Number Fields  
   5.1 Definitions and Bounds  
   5.2 Lattice Basis Reduction and Reconstructing Algebraic Numbers  
   5.3 Ratios of Radicals in Algebraic Number Fields  
   5.4 Approximating Radicals and Ratios of Radicals  
   5.5 A Probabilistic Test for Equality  
   5.6 Sums of Radicals over Algebraic Number Fields  
6 Denesting Radicals - The Basic Results  
   6.1 The Basic Theorems  
   6.2 Denesting Sets and Reduction to Simple Radical Extensions  
   6.3 Characterizing Denesting Elements  
   6.4 Characterizing Admissible Sequences  
   6.5 Denesting Radicals - The Algorithms  
7 Denesting Radicals - The Analysis  
   7.1 Preliminaries  
   7.2 Description and Analysis of Step 1  
   7.3 Description and Analysis of Step 2  
   7.4 Description and Analysis of Step 3  
   7.5 The Final Results  
Appendix: Roots of Unity in Radical Extensions of the Rational Numbers  
Summary
<table>
<thead>
<tr>
<th>Zusammenfassung</th>
<th>189</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lebenslauf</td>
<td>191</td>
</tr>
</tbody>
</table>
1 Introduction

Problems and Results An important issue in symbolic computation is the simplification of expressions. Since many algorithms in Computer Algebra systems like Mathematica, Maple, and Reduce work in quite general settings they do not necessarily find a solution to a given problem described in the easiest possible way. Simplification algorithms can be applied to express these solutions in a form that is more convenient for later use. For example, to determine whether the solution itself or the difference of two such solutions is zero.

In this thesis we consider simplification algorithms for radical expressions. A radical over a field $F$ is a root $\sqrt[d]{\rho}$ of some element $\rho$ in the field $F$. We prefer the name radical since there are $d$ different roots of $\rho$ and in general we may not be referring to an arbitrary $d$-th root of $\rho$ but to a certain fixed value of $\sqrt[d]{\rho}$. A radical expression over $F$ is any expression built from the usual arithmetic operations and from, possibly nested, roots.

Dealing with radicals has a long history in mathematics. For example, Galois Theory emerged from the problem of solving polynomials by radicals. It seems that in Computer Science people first got interested in radicals by their connection to the bit complexity of certain optimization problems such as the Traveling Salesman Problem (TSP) or the Shortest Path Problem (SPP). In fact, in any solution to these problems the length of tours or paths have to be compared. Assuming that the cities in the TSP have rational coordinates then the length of a path is given by a sum of square roots of rational numbers. Therefore in order to compare the length of two tours the sign of a sum of square roots has to be determined.

This quite innocent looking problem turned out to be extremely difficult and until now no efficient solution, not even a promising approach, is known. By efficient we mean an algorithm that determines the sign by a number of bit operations that is polynomial in the length of the sum and in the bit size of the rational numbers.

It was exactly this problem that stimulated our interest in the question we answer in the first part of this thesis. We show that although the sign of a sum of square roots may not be computable in polynomial time it is nevertheless possible to decide in polynomial time whether a sum of square roots is zero. Perhaps somewhat surprising the solution turns out to be quite simple. Moreover, with some effort we extend the result to sums of real radicals of arbitrary degree and generalize the solution by replacing the field $\mathbb{Q}$ by real algebraic number fields.
These fields are easily defined as follows. Let \( \alpha \) be the root of a polynomial with rational coefficients. The smallest field containing \( \mathbb{Q} \) and \( \alpha \) is called the algebraic number field generated by \( \alpha \) and is denoted by \( \mathbb{Q}(\alpha) \). Since exact arithmetic is possible in these fields they play an important role in symbolic computation (see for example [Lo1]).

To determine whether a sum of radicals over the field \( \mathbb{Q} \), say, is zero it is first simplified into a sum such that the radicals appearing in the transformed sum are linearly independent. Then it is easy to check whether the original sum is zero. This can happen if and only if the coefficients in the sum resulting from the transformation step are zero.

The algorithms that transforms an arbitrary sum \( S \) of real radicals into a sum of linearly independent radicals is based on the following result due to C. L. Siegel [Si].

Let \( F \) be any real field, i.e., \( F \subseteq \mathbb{R} \). If for positive integers \( d_i \) and elements \( \rho_i \in F, i = 1, \ldots, k \), \( \sqrt[d_i]{\rho_i} \in \mathbb{R} \) and

\[
\frac{\sqrt[d_1]{\rho_1}}{\sqrt[d_j]{\rho_j}} \notin F \text{ for all } i \neq j,
\]

then \( \sqrt[d_1]{\rho_1}, \ldots, \sqrt[d_k]{\rho_k} \) are linearly independent over \( F \).

Hence we easily do the transformation by determining for any pair of radicals in \( S \) whether its ratio is in \( F \), by computing the representation of the ratio in \( F \), and by finally collecting terms in order to determine the coefficients.

If we were satisfied with an algorithm whose run time is polynomial in the degrees \( d_i \) rather than in \( \log d_i \) the ratio test is simple. But to reduce the run time to a polynomial in \( \log d_i \) even in the case \( F = \mathbb{Q} \) we have to work much harder. For algebraic number fields we achieve this reduction only by allowing the algorithm to give with small probability an incorrect answer, that is, by an algorithm of Monte-Carlo-type.

We also show how to determine for sums of complex radicals over certain complex algebraic number fields whether they are zero. In that case, however, the run times are polynomial only in the degrees of the radicals themselves.

The result of Siegel mentioned above has some history. A special case of it was first proven in 1940 by A. S. Besicovitch [Be]. This result was slightly generalized in 1953 by L. J. Mordell until in 1971 Siegel proved the theorem in the form stated above. But this is still not the end of the story because in 1974 M. Kneser generalized Siegel’s result to certain complex fields. In this
thesis we also contribute a bit to this history by giving a fairly short and simple proof of Siegel's Theorem. Moreover, we show how it can be used to prove and generalize certain results from Kummer Theory.

By generalizing the results above to sums of radicals that are not necessarily defined over the same field we naturally hit upon the problem of simplifying or denesting nested radicals.

Throughout the last years this problem has been studied intensively by various mathematicians and computer scientists ([BFHT],[HH],[La2],[La3],[Z]). Many of them have been attracted by equations of Ramanujan [R] such as the following:

\[
\sqrt[3]{\sqrt{2} - 1} = \frac{\sqrt[3]{1}}{9} - \frac{\sqrt[3]{2}}{9} + \frac{\sqrt[3]{4}}{9}
\]

\[
\sqrt[3]{\sqrt{5} - \sqrt{4}} = \frac{1}{3} \left( \sqrt{2} + \sqrt[3]{20} - \sqrt[3]{25} \right)
\]

\[
\sqrt[6]{7 \sqrt{20} - 19} = \frac{\sqrt[3]{5}}{3} - \sqrt[3]{3}
\]

These examples sufficiently explain the problem itself. Denesting a radical expression means decreasing its nesting depth over a field \( F \). For example, the depth over \( \mathbb{Q} \) of the left-hand side of Ramanujan’s equations is 2 while the depth of the right-hand side is 1.

If we are given a nested radical and are asked to denest it then this is at first not a meaningful question because for different values of the roots involved we may get different denestings. For example, if \( \sqrt{\sqrt{5} + 2} - \sqrt{\sqrt{5} - 2} \) has to be denested, in general, we will assume that the two square roots have the same value and that perhaps the real third roots are meant. In this case the formula denests to 1. But if one of the square roots is positive and the other one negative then the formula denests to a square root of 5 provided we still take real third roots. And for the complex third roots the denestings are again different. So the values of the roots involved in the nested radical have to be specified in advance. For example, in case of real radicals of even degree we will assume in this thesis that their value is given by the positive real root.

In the last years a lot of progress has been made for the problem of denesting radicals. Landau [La2] showed how to compute a denesting for a nested radical whose nesting depth is just one off the optimal one. Horng and Huang [HH] achieved a minimal denesting and also showed how to solve a polynomial by a radical of minimum nesting depth, if it is solvable.
by radicals at all. However, in both cases the denestings are with respect to a field containing all roots of unity. As it turns out for any radical expression a single root of unity can be determined such that a denesting over the field generated by this root of unity is (almost in case of [La2]) a minimum depth denesting of the expression over the field containing all roots of unity. The degree of this root, however, is either single- (Landau) or even double-exponential (Horng, Huang) in the description size of the minimal polynomials of the expressions to be denested.

In particular, these results cannot explain or compute the special form of Ramanujan’s examples above since in these examples no roots of unity are required. But even if they did compute these denestings they would do so very inefficiently. Moreover, it looks like the approach of Landau and Horng, Huang cannot yield efficient algorithms. They use Galois Theory and have to compute the splitting field of the radical expression they want to denest, that is, the field generated by all roots of the minimal polynomial of the radical. The degree of this extension may be exponential in the degree of the roots appearing in the radical expression.

In this thesis we follow a different approach that leads to several efficient denesting algorithms although these algorithms produce minimum depth denestings only of a restricted kind. Therefore they are not really comparable to the results of Landau and Horng, Huang. On the other hand, our theoretical results do explain Ramanujan’s denestings and our algorithms efficiently determine them. Moreover, several of our results indicate that there is good reason to believe that the methods of this thesis may even lead to more efficient algorithms for minimum depth denestings in general.

To be more specific, we prove for example that if $\gamma$ is an element of some field $F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}), F \subset \mathbb{R}$, $q_i \in F$, $\sqrt[d_i]{\rho_i} \in \mathbb{R}$, and if $\sqrt[2]{\gamma}$ is a real nested radical of depth 2 then $\sqrt[2]{\gamma}$ can be written as a depth 1 expression over $F$ if and only if there exists an element $\gamma_0 \in F$ such that

$$\sqrt[d_N]{\gamma_0} \sqrt[2]{\gamma} \in F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}).$$

$N$ is the degree of the extension $F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}) : F$. This theorem is a generalization of a similar result due to Borodin et al. [BFHT] that is restricted to certain expressions of square roots.

We then show how to compute the element $\gamma_0$ efficiently. This problem was considered before in [Z] and [La3] although the authors did not know the theorem above and hence were not aware of the fact that elements $\gamma_0$ as in the theorem lead to the only possible form of a denesting. Basically we improve
the algorithms in these two papers from exponential to polynomial time and achieve the first denesting algorithms whose run times are polynomial in the maximal output size.

Before we finish this brief description of the results obtained in this thesis let us mention the basic technique of our algorithms. Using a variant of the Kannan, Lenstra, and Lovász algorithm to (re-)construct minimal polynomials we show how to compute in polynomial time the exact representation of an element \( \beta \) in an algebraic number field \( \mathbb{Q}(\alpha) \), provided an upper bound \( B \) on the representation size of \( \beta \) is given and an approximation \( \bar{\beta} \) to \( \beta \) is known such that the number of correct bits in \( \bar{\beta} \) is roughly quadratic in \( B \). We believe that this technique will have a lot of applications in algorithmic algebra and number theory.

**The Model of Computation** Throughout this thesis we are only interested in the *bit complexity* of problems. But this does not uniquely define a model of computation. We can assume for example a (multi-tape) Turing Machine as it is defined in any standard textbook on algorithm theory like [AHU],[CLR]. But we can also use the following variant of a Random Access Machine (RAM) as it is defined in [Sc2].

As usual the RAM has a storage consisting of an infinite number of locations indexed with the positive integers. Each location can store an arbitrary integer. Furthermore the RAM has a so-called processing unit.

The operations allowed consist of the usual storage operations like “load” and “store”. But the only arithmetic operation allowed is the successor function. Finally comparisons are allowed. The time needed for each operation is defined as the logarithm \( \log \) of the maximum bit size of the operands. For further details we refer to [Sc2].

Yet another model of computation are the Pointer Machines which are also defined precisely in [Sc2].

In these models the best currently known upper bounds for multiplying two \( n \)-bit integers are pairwise distinct. They are \( O(n \log n \log \log n) \), \( O(n \log n) \), and \( O(n) \), respectively. To make our results easily applicable to these and other models of computation we chose to represent the run times in terms of the number of what we call *elementary operations* and in terms of the *maximum bit size* of the numbers on which to perform these operations.

\[ \sqrt[n]{x} \]

In this thesis \( \log \) always denotes the logarithm to the base 2; the logarithm to the base \( e \) is denoted by \( \ln \).
We define precisely what numbers we allow and which operations are considered as elementary.

First, we consider integers $z$ represented in binary. We call $z$ an $n$-bit integer if its binary length is $n$. Elementary operations on integers are additions, subtractions, multiplications, and divisions with remainder.

But we also consider floating-point numbers. Real floating-point numbers are represented by a pair $(b, m) \in \mathbb{Z} \times \mathbb{Z}$ which represents the number $b2^m$. $(b, m)$ is called an $n$-bit floating-point number if the sum of the bit size of $b$ and of the bit size of $m$ is $n$. Complex floating-point numbers are represented by a pair of real floating-point numbers, the first being the real part, the second one being the imaginary part. An $n$-bit complex floating number is a floating-point number such that the sum of the representation size of the real and imaginary part is $n$.

As for the integers elementary operations on floating-point numbers are additions, subtractions, and multiplications. But also computing the inverse or the square root of an $n$-bit floating-point number with relative error $O(2^{-n})$ are considered as elementary.

As is well-known in the three models mentioned above, if $M(n)$ denotes the time needed to multiply two $n$-bit integers then any elementary operation on integers or floating-point numbers can be done in $O(M(n))$ time. This may justify that we treat these operations in the same way.

Our main results show that certain simplification problems on radicals can be solved by a polynomial number of elementary operations on integers and floating-point numbers of polynomial size. This implies that these problems can be solved in polynomial time in any reasonable model of computation that allows bit operations. The exact run times for a specific model of computation can be deduced from our general results by plugging in the maximum number of bit operations needed for an elementary operation in this model of computation.

The elementary operations on floating-point numbers clearly include the elementary operations on integers. Therefore the distinction between elementary operations on integers and on floating-point numbers deserves some explanation. On the one hand, the algorithms we are going to describe in this thesis return as results elements in algebraic number fields represented as a linear combination of some basis elements with rational coefficients. These coefficients have to be represented exactly. But by any definition of floating-point numbers some rational numbers cannot be represented exactly by a single floating-point number. For example, in the system defined above $\frac{1}{3}$ is not representable exactly. Describing the coefficients as tuples of
floating-point numbers does not seem to be an appropriate form. At least it
does not coincide with the usual understanding of a floating-point number.

Finally we want to be able to perform exact arithmetic in algebraic num-
ber fields. This is best described by exact arithmetic on integers. For exam-
ple, these operations often require greatest common divisor computations.
Describing these computations by elementary operations on floating-point
numbers is a crude abuse of language and notation.

On the other hand, based on approximation algorithms due to Schönhage
and Brent we describe approximation algorithms for algebraic numbers. To
describe these approximation algorithms via elementary operations on inte-
gers is at least inaccurate and confusing. Although as far a asymptotic run
times are concerned and one is very careful it would not cause too much
trouble. In particular, we are never forced to represent numbers by floating-
point approximations because they are either too large or too small.

We also believe that by distinguishing between elementary operations on
integers and elementary operations on floating-point numbers the analysis
of the algorithms described in this thesis are easier to understand. Recall for
example from the previous paragraph that our basic technique transforms
an approximation to an algebraic number given by a floating-point number
into an exact representation as a linear combination of certain basis elements
of an algebraic number field.

Finally let us mention that division and taking square roots have been
included into the elementary operations on floating-point numbers since in
the three models mentioned in the beginning it is correct that the time
needed for these procedures is asymptotically the same as the one needed
for multiplication. However, it is not correct that division and taking square
roots can be done by a constant number of arithmetic operations on inte-
gers of size $O(n)$. Since the approximation algorithms of Brent [Br] and
Schönhage [Sc3] apply divisions and square rootings we cannot accurately
state the run times of our approximation algorithms in terms of arithmetic
operations only$^2$.

For the algorithms we have to determine certain constants, for example,
if approximations with certain precision are required. In general, the con-
stants we derive will not be optimal. Deriving better constants would often
complicate the notation and add more technical details that are not crucial
to our algorithms.

$^2$In particular, Brent’s algorithms use the Arithmetic-Geometric-Mean-Iteration.
For run times we use of course the standard $O$-notation.

Our main interest is to show that certain simplification problems for radicals are solvable in polynomial time. Although the run times we deduce are asymptotically the best we can prove so far, for many subalgorithms they are not optimal. Instead the analysis of these subalgorithms only shows that their run times will not determine the overall run time. The reader should always be aware of this fact.

A brief overview In Section 2 we present all the material from algebra and number theory that will be used in the thesis. It contains a brief description of the basics from field theory and summarizes the main results from Galois Theory. We also define algebraic numbers and algebraic integers and mention their main properties. In Section 3 we give a short and fairly simple proof of Siegel’s Theorem. In Section 4 by restricting ourselves to the rational numbers we demonstrate the basic ideas of the algorithm that transforms a sum of radicals into a sum of linearly independent radicals. In Section 5 we generalize the algorithm to algebraic number fields. In particular, the reconstruction algorithm for the exact representation of algebraic numbers is described.

With Section 6 the part on denesting radicals begins. In Section 6 itself we prove all the relevant theoretical facts and outline the basic algorithms that compute the denestings. For example, if $\gamma$ is a sum of radicals over $F$ we describe an algorithm that determines whether an element $\gamma_0 \in F$ exists such that $\sqrt[\gamma_0]{} \sqrt[\gamma]{\gamma}$ can be written as a sum of radicals over $F$. If such an element exists the algorithm computes one. In Section 7 we fill in the details of the algorithms and give a precise analysis.

Acknowledgement Without the help and encouragement of Helmut Alt, Susan Landau, Emo Welzl, and Chee Yap this thesis would not have been written.

My advisor, Helmut Alt, gave me the opportunity to work in a field that does not belong to his own primary interests. He was an extremely careful reader of the thesis. He not only found a lot of confusing mistakes in the first versions but also tried to improve my style. He, and Emo Welzl, also gave very useful hints. Helmut Alt drew my attention to the approximation algorithms used in the thesis and Emo Welzl suggested various methods for the probabilistic algorithm of Section 5. Both of them permanently encouraged me to try again if first attempts to solve a problem failed and
I was convinced that an efficient solution did not exist or that I would not find it.

Susan Landau pointed out to me the problem of denesting radicals and suggested that my previous techniques might be useful for this problem. Without her encouragement the second part of the thesis would not have been written.

Had Chee Yap not spent the academic year 1989/90 at the Free University Berlin this thesis would not have been written. In his course on Computer Algebra I first learned about many bounds used in the thesis and, most important, I learned about lattice reduction. His class notes were an invaluable help when working out the technical details of the algorithms. Since his forthcoming book on Computer Algebra [Y] is not yet published I decided to cite the original research papers for bounds and basic techniques. But most of this material will be contained in Chee Yap’s book. But Chee Yap not only introduced me to Computer Algebra. He also introduced me to the problem that started this thesis. After trying in vain for months to come up with an efficient algorithm to determine the sign of a sum of square roots of integers, he finally suggested to see whether we could at least come up with an algorithm that checks whether a sum of square roots is zero. The first solution to this problem (using only the Primitive Element Theorem) was joint work with Chee Yap. He also suggested to consider not only square roots but arbitrary radicals. This eventually lead to the results of this thesis.
2 Basics from Algebra and Algebraic Number Theory

In this section we review the basic facts from algebra and algebraic number theory which will be used in this thesis. This material can be found in any textbook on algebra and algebraic number theory like [J], [Ja],[I], or [M]. So most of the proofs are omitted. Furthermore we try to keep the exposition as simple and self-contained as possible. As a consequence the reader familiar with algebra and number theory will find that a lot of facts mentioned can be generalized considerably.

The most fundamental concept used is that of an algebraic element. Consider an arbitrary subfield \( F \) of the complex numbers \( \mathbb{C} \). An element \( \alpha \in \mathbb{C} \) is called algebraic over \( F \) if a polynomial \( p(X) = \sum_{i=0}^{n} p_i X^i \), \( p_i \in F \), exists such that \( p(\alpha) = 0 \). An important example is the case \( F = \mathbb{Q} \). If \( \alpha \in \mathbb{C} \) is algebraic over \( \mathbb{Q} \) we simply call \( \alpha \) algebraic. It is not very hard to show that the algebraic numbers in \( \mathbb{C} \) over \( F \) form a field with respect to addition and multiplication in \( \mathbb{C} \).

If \( \alpha \in \mathbb{C} \) is algebraic over \( F \), then the smallest degree polynomial \( p(X) \in F[X] \) with \( p(\alpha) = 0 \) and leading coefficient \( p_n = 1 \) is called the minimal polynomial of \( \alpha \) over \( F \). Hence the minimal polynomial is an irreducible element of \( F[X] \). The degree of the minimal polynomial is also called the degree of \( \alpha \) over \( F \).

For example, all rational numbers \( q \) are algebraic over \( \mathbb{Q} \) with minimal polynomial \( X - q \). Furthermore \( \sqrt[d]{q} \), \( q \in \mathbb{Q} \) and \( d \in \mathbb{N} \), is algebraic and the minimal polynomial of \( \sqrt[d]{q} \) is a divisor of \( X^d - q \).

A field \( E \supseteq F \) is called an algebraic extension of \( F \) if all elements of \( E \) are algebraic over \( F \). Such an extension will be denoted by \( E : F \). \( E \) may be considered as a vector space over \( F \). If this vector space has finite dimension \( n \) then \( n \) is called the degree of the extension and is denoted by \( [E : F] \). Any vector space basis is called a basis of the field extension or simply a field basis.

Important examples of algebraic extensions are extensions generated by adjoining a single algebraic number to a field \( F \). To define this more precisely let \( \alpha \in \mathbb{C} \) be algebraic over \( F \) and assume that the minimal polynomial \( p(X) \) of \( \alpha \) over \( F \) is of degree \( n \) (denoted by \( \deg p = n \)). Then the smallest field containing \( F \) and \( \alpha \) is denoted by \( F(\alpha) \). Since this field is isomorphic to the field of all polynomials in \( F[X] \) taken modulo the minimal polynomial \( p \).
of \( \alpha \) (denoted \( F[X]/(p) \)) it is easily seen that \( F(\alpha) = \{ \sum_{i=0}^{n-1} \kappa_i \alpha^i | \kappa_i \in F \} \).

Moreover, the degree of \( F(\alpha) \) over \( F \) is \( n \), the degree of the minimal polynomial of \( \alpha \) over \( F \).

This can be generalized in the following way. Consider a field \( F \) and \( k \) numbers \( \alpha_1, \ldots, \alpha_k \) that are algebraic over \( F \). By \( F(\alpha_1, \ldots, \alpha_k) \) denote the smallest field containing \( F \) and the numbers \( \alpha_1, \ldots, \alpha_k \). Also let \( n_i \) be the degree of the minimal polynomial \( p_i \) of \( \alpha_i \) over \( F(\alpha_1, \ldots, \alpha_{i-1}) \). Then the degree \( n \) of the extension is \( n = \prod_{i=1}^k n_i \) and a basis is given by the elements

\[
\alpha_1^{e_1} \alpha_2^{e_2} \cdots \alpha_k^{e_k}, \quad 0 \leq e_1 < n_1, 0 \leq e_2 < n_2, \ldots, 0 \leq e_k < n_k.
\]

So any element in \( F(\alpha_1, \ldots, \alpha_k) \) can uniquely be written as

\[
\sum_{e_1=0}^{n_1-1} \sum_{e_2=0}^{n_2-1} \cdots \sum_{e_k=0}^{n_k-1} \kappa_{e_1,e_2,\ldots,e_k} \alpha_1^{e_1} \alpha_2^{e_2} \cdots \alpha_k^{e_k}
\]

with \( \kappa_{e_1,e_2,\ldots,e_k} \in F \).

We are interested in the isomorphisms of \( F(\alpha_1, \alpha_2, \ldots, \alpha_k) \) into subfields of \( \mathbb{C} \) that fix \( F \) pointwise. These mappings are called the embeddings of \( F(\alpha_1, \alpha_2, \ldots, \alpha_k) \) into \( \mathbb{C} \). Furthermore the images of \( F(\alpha_1, \alpha_2, \ldots, \alpha_k) \) under the isomorphisms are called the conjugate fields of \( F(\alpha_1, \alpha_2, \ldots, \alpha_k) \).

First let us assume that the extension is generated by a single element \( \alpha \). Let \( p \) be the minimal polynomial of \( \alpha \) over \( F \), \( \deg p = n \). Since we assume \( \mathbb{Q} \subseteq F \) and \( p \) is irreducible, \( p \) has exactly \( n \) distinct roots \( \alpha = \alpha^{(0)}, \alpha^{(1)}, \ldots, \alpha^{(n-1)} \), called the conjugates of \( \alpha \).

Since \( \sigma(p(\alpha)) = p(\sigma(\alpha)) \) any embedding \( \sigma \) of \( F(\alpha) \) must map \( \alpha \) onto one of its conjugates. Hence there are at most \( n \) distinct embeddings. On the other hand, each field \( F(\alpha^{(i)}), i = 0, 1, \ldots, n-1 \), is isomorphic to the field \( F[X]/(p) \). Therefore there are exactly \( n \) distinct embeddings of \( F(\alpha) \) which are given by

\[
\sigma_i : F(\alpha) \longrightarrow F(\alpha^{(i)}) \quad \sum_{j=0}^{n-1} \kappa_j \alpha^j \longrightarrow \sum_{j=0}^{n-1} \kappa_j \alpha^{(i)j}.
\]

This can be generalized in the following way to field extensions \( F(\alpha_1, \ldots, \alpha_k) \) generated by more than one algebraic element. Assume the embeddings of \( F(\alpha_1, \ldots, \alpha_{k-1}) \) have already been determined. Let \( \tau \) be such

\footnote{For \( i = 1 \) this is \( F \). We apply a similar convention in many situations below.}
an embedding. If \( p_k(X) = \sum_{j=0}^{n_k} \mu_j X^j \), \( \mu_j \in F(\alpha_1, \ldots, \alpha_{k-1}) \), is the minimal polynomial of \( \alpha_k \) over \( F(\alpha_1, \ldots, \alpha_{k-1}) \), denote by \( \tau(p_k) \) the polynomial \( \sum_{j=0}^{n_k} \tau(\mu_j) X^j \). Since \( p_k \) is irreducible over the field \( F(\alpha_1, \ldots, \alpha_{k-1}) \) so is \( \tau(p_k) \) over \( F(\tau(\alpha_1), \ldots, \tau(\alpha_{k-1})) \).

We can extend \( \tau \) by mapping \( \alpha_k \) onto any of the \( n_k \) distinct roots of \( \tau(p_k) \). Since \( F(\alpha_1, \ldots, \alpha_k) \) is isomorphic to the field of polynomials in \( F(\alpha_1, \ldots, \alpha_{k-1})[X] \) taken modulo \( p_k \) it is easily verified that these extensions are indeed embeddings of \( F(\alpha_1, \ldots, \alpha_k) \).

On the other hand, if \( \sigma \) is an embedding of \( F(\alpha_1, \ldots, \alpha_k) \) over \( F \) then its restriction \( \tau \) to \( F(\alpha_1, \ldots, \alpha_{k-1}) \) must be an embedding of \( F(\alpha_1, \ldots, \alpha_{k-1}) \). Since

\[
0 = \sigma(p_k(\alpha_k)) = \tau(p_k)(\sigma(\alpha_k))
\]

it follows that any embedding of \( F(\alpha_1, \ldots, \alpha_k) \) must have the form described above.

Summarizing the following theorem has been shown.

**Theorem 2.1** Let \( F \subset C \) be a field and \( \alpha_1, \alpha_2, \ldots, \alpha_k \) algebraic over \( F \). If \( [F(\alpha_1, \ldots, \alpha_i) : F(\alpha_1, \ldots, \alpha_{i-1})] = n_i \) and \( n = [F(\alpha_1, \ldots, \alpha_k) : F] = \prod_{i=1}^{k} n_i \), then \( F(\alpha_1, \ldots, \alpha_k) \) has exactly \( n \) distinct embeddings over \( F \). Moreover, any embedding of \( F(\alpha_1, \ldots, \alpha_{i-1}) \) over \( F \) can be extended in exactly \( n_i \) different ways to an embedding of \( F(\alpha_1, \ldots, \alpha_i) \) over \( F \).

Furthermore note that if \( P_i \) is the minimal polynomial of \( \alpha_i \) over \( F \) then any conjugate field of \( F(\alpha_1, \ldots, \alpha_k) \) has the form \( F(\alpha_{i,j_1}, \ldots, \alpha_{k,j_k}) \), where \( \alpha_{i,j_i} \) is a root of \( P_i \). In fact, for any embedding \( \sigma \) of \( F(\alpha_1, \ldots, \alpha_k) \) over \( F \)

\[
0 = \sigma(P_i(\alpha_i)) = P_i(\sigma(\alpha_i)).
\]

Finally let us mention that if \( \gamma \) is an element of \( F(\alpha_1, \ldots, \alpha_k) \) then the degree \( m \) of the minimal polynomial of \( \gamma \) must divide \( n \), since \( F(\gamma) \) is a subspace of the vector space \( F(\alpha_1, \ldots, \alpha_k) \). Any embedding \( \sigma \) of \( F(\alpha_1, \ldots, \alpha_k) \) over \( F \) must map \( \gamma \) onto one of its conjugates over \( F \) and for each conjugate \( \gamma' \) of \( \gamma \) there are exactly \( \frac{n}{m} \) embeddings \( \sigma \) that map \( \gamma \) onto \( \gamma' \).

Important extensions are those generated by the roots \( \alpha_0, \alpha_1, \ldots, \alpha_{n-1} \) of a single polynomial \( p \) of degree \( n \). These fields are called splitting fields or **Galois extensions**. A Galois extension coincides with all its conjugate fields,

\[\text{Observe that unless } P_i = p_i, p_i \text{ the minimal polynomial of } \alpha_i \text{ over } F(\alpha_1, \ldots, \alpha_{i-1}), \text{ not all combinations of roots of } P_1, \ldots, P_k \text{ are possible.}\]

\[\text{This definition is only correct since we are restricting ourselves to subfields of the complex numbers.}\]
hence the embeddings of such an extension are not only isomorphisms but automorphisms. Moreover, these automorphisms form a group called the Galois group of the extension. The Fundamental Theorem of Galois Theory describes a one-to-one correspondence between the subgroups of the Galois group of a Galois extension and the subfields of the extension.

If \( F \subset L \), \( L \) a subfield of the Galois extension \( E = F(\alpha_0, \alpha_1, \ldots, \alpha_{n-1}) \) over \( F \), and \( \sigma \) is an element of the Galois group \( G \) of \( E \) then we say that \( \sigma \) fixes \( L \) if and only if
\[
\sigma(\gamma) = \gamma, \text{ for all } \gamma \in L.
\]
The set of all \( \sigma \in G \) that fix \( L \) form a subgroup of \( G \), we denote this group by \( \text{Gal}(L) \). Hence \( \text{Gal} \) is a function that maps a subfield \( L \) of \( E \) onto a subgroup of \( G \). \( E \) can obviously be considered as a Galois extension of any subfield \( L \). As it turns out the Galois group of \( E \) over \( L \) is \( \text{Gal}(L) \).

On the other hand, for any subgroup \( H \) of \( G \) we consider the set of elements \( \gamma \in E \) such that
\[
\sigma(\gamma) = \gamma, \text{ for all } \sigma \in H.
\]
It is easily seen that for any subgroup \( H \) these elements form a subfield of \( E \). This field is called the fixed field of \( H \) and is denoted by \( \text{Inv}(H) \).

**Theorem 2.2 (Fundamental Theorem of Galois Theory)** Let \( E \) be a Galois extension of \( F \subset C \) with Galois group \( G \). By \( \Gamma \) denote the set of all subgroups \( H \) of \( G \) and by \( \Sigma \) denote the set of all subfields \( L \) of \( E \) with \( F \subset L \). The mappings
\[
H \mapsto \text{Inv}(H)
\]
\[
L \mapsto \text{Gal}(L)
\]
are inverses and hence bijective mappings of \( \Gamma \) onto \( \Sigma \) and from \( \Sigma \) onto \( \Gamma \). Moreover, the extension \( L : F \) is a Galois extension if and only if \( \text{Gal}(L) \) is a normal subgroup of \( G \). In this case the Galois group of \( L \) over \( F \) is isomorphic to \( G/\text{Gal}(L) \).

Any textbook on algebra contains a proof of this theorem.

Important examples of Galois extensions are the so-called cyclotomic fields. Consider the equation
\[
X^n - 1 = 0.
\]
Any root of this equation is called an \textit{n-th root of unity}. These roots form a multiplicative group and, since this group is a finite subgroup of \( C \setminus \{0\} \), this group is cyclic (Any finite subgroup of the multiplicative group of a field \( F \) is cyclic [J]). Therefore a root \( \zeta_n \) of this equation exists such that \( n \) is the smallest integer with \( \zeta_n^n = 1 \). Any root of \( X^n - 1 = 0 \) with this property is called a \textit{primitive n-th root of unity}. In particular, if \( \zeta_n \) is a primitive \( n \)-th root of unity then any root of \( X^n - 1 = 0 \) is a power of \( \zeta_n \). Hence the \textit{n-th cyclotomic field} \( \mathbb{Q}(\zeta_n) \) is independent of the choice of the primitive root \( \zeta_n \) and is a Galois extension of \( \mathbb{Q} \).

As is well-known the degree of the extension \( \mathbb{Q}(\zeta_n) : \mathbb{Q} \) is \( \varphi(n) \), where \( \varphi \) denotes Euler’s \( \varphi \)-function counting the number of integers between 1 and \( n \) relatively prime to \( n \). Moreover, the Galois group of \( \mathbb{Q}(\zeta_n) : \mathbb{Q} \) is isomorphic to the multiplicative group \( \mathbb{Z}_n^* \) of integers between 1 and \( n \) taken modulo \( n \) which are relatively prime to \( n \). Since this group is abelian the Fundamental Theorem of Galois Theory implies that any subfield \( F, \mathbb{Q} \subset F \subset \mathbb{Q}(\zeta_n) \), is a Galois extension of \( \mathbb{Q} \). We summarize these facts in

**Lemma 2.3** Let \( \zeta_n \) be a primitive \( n \)-th root of unity. The \textit{n-th cyclotomic field} \( \mathbb{Q}(\zeta_n) \) is a Galois extension of \( \mathbb{Q} \). Furthermore the Galois group of this extension is isomorphic to the group \( \mathbb{Z}_n^* \). Hence the Galois group is abelian and all subfields of \( \mathbb{Q}(\zeta_n) \) are Galois extensions of \( \mathbb{Q} \).

If \( E \) is an arbitrary extension of \( \mathbb{Q} \) then \( E(\zeta_n) \) is a Galois extension of \( E \) for any root of unity \( \zeta_n \). The next theorem relates the Galois group of \( E(\zeta_n) \) over \( E \) to the Galois group of \( \mathbb{Q}(\zeta_n) \) over a certain subfield of \( \mathbb{Q}(\zeta_n) \).

**Theorem 2.4** Let \( E \) be a Galois extension of the field \( K \). Denote the Galois group of this extension by \( G \). Assume furthermore that \( F \) is an arbitrary extension of \( K \) and denote by \( EF \) the smallest field containing \( E \) and \( F \). Then the field \( EF \) is a Galois extension of \( F \) and the Galois group of \( EF : F \) is isomorphic to the subgroup of \( G \) corresponding to the extension \( E : F \cap E \).

A proof of this theorem can be found in [L].

Hence for any extension \( E \) of \( \mathbb{Q} \) the Galois group of \( E(\zeta_n) \) over \( E \) is isomorphic to the Galois group of \( \mathbb{Q}(\zeta_n) \) over \( \mathbb{Q}(\zeta_n) \cap E \). In particular, it is abelian.

So far extensions of the form \( F(\alpha_1, \ldots, \alpha_k) \) have been considered but by the well-known \textit{Primitive Element Theorem} any extension of \( F \subset \mathbb{C} \) generated by a finite number of algebraic numbers over \( F \) can already be generated by adjoining a single algebraic number to \( F \). Using the results mentioned so far we derive a quantitative version of this theorem.
**Lemma 2.5** Let $F \subset \mathbb{C}$ be a field and $E = F(\alpha_1, \ldots, \alpha_k)$ an algebraic extension of $F$ with $[E : F] = n$. Denote the different embeddings of $E$ over $F$ by $\sigma_0, \sigma_1, \ldots, \sigma_{n-1}$.

Assume $\gamma \in E$ such that

$$\sigma_i(\gamma) \neq \sigma_j(\gamma), \text{ for all } i \neq j,$$

then $E = F(\gamma)$.

**Proof:** Consider for each embedding $\sigma_i$ its restriction $\tau_i$ to $F(\gamma)$. $\tau_i$ is an embedding of $F(\gamma)$ over $F$. Since $\sigma_i(\gamma) \neq \sigma_j(\gamma)$ the restrictions $\tau_i$ are distinct embeddings of $F(\gamma)$. Hence $F(\gamma)$ has at least $n$ distinct embeddings over $F$ and therefore $[F(\gamma) : F] \geq n$. On the other hand, $F(\gamma)$ is a subfield of $E$. So its degree must be less than $n$. This shows $n = [F(\gamma) : F]$. But no field extension of degree $n$ has a proper subfield of degree $n$. This finally proves $F(\gamma) = E$. 

As an immediate consequence from this lemma we get the *Primitive Element Theorem* which we state in its classical form.

**Theorem 2.6 (Primitive Element Theorem)** Let $F$ be a field that contains the rational numbers. Assume that $\alpha, \beta$ are algebraic over $F$. Furthermore denote by $\alpha_0 = \alpha, \alpha_1, \ldots, \alpha_{n-1}$ and $\beta_0 = \beta, \beta_1, \ldots, \beta_{m-1}$ the conjugates of $\alpha$ over $F$ and of $\beta$ over $F$, respectively. If $\kappa \in F$ satisfies

$$\alpha_i + \kappa \beta_k \neq \alpha_j + \kappa \beta_l$$

for all conjugates $\alpha_i, \alpha_j$ of $\alpha$ and all conjugates $\beta_k, \beta_l$ of $\beta$ such that $i \neq j$ or $k \neq l$ then

$$F(\alpha, \beta) = F(\alpha + \kappa \beta).$$

**Proof:** Consider two different field embeddings $\sigma, \tau$ of $F(\alpha, \beta)$. $\sigma(\alpha) = \alpha_i$ and $\sigma(\beta) = \beta_k$ and, likewise, $\tau(\alpha) = \alpha_j$, $\tau(\beta) = \beta_l$ for conjugates $\alpha_i, \alpha_j$ of $\alpha$ and conjugates $\beta_k, \beta_l$ of $\beta$. Since any embedding is completely determined by its action on $\alpha$ and $\beta$ the two embeddings can be different if and only if $\alpha_i \neq \alpha_j$ or $\beta_k \neq \beta_l$. But then

$$\sigma(\alpha + \kappa \beta) = \alpha_i + \kappa \beta_k \neq \alpha_j + \kappa \beta_l = \tau(\alpha + \kappa \beta).$$

The theorem follows from Lemma 2.5.
Since $F$ contains infinitely many elements and the condition of the theorem excludes at most $\frac{n(n-1)}{2} \frac{m(m-1)}{2}$ elements from $F$ the theorem actually states that any extension of $F$ generated by two elements, and by induction on the number of generators that any extension generated by a finite number of elements, can already be generated by a single algebraic number over $F$.

We nevertheless consider algebraic extensions of the form $F(\alpha_1, \ldots, \alpha_k)$ because in going from $F(\alpha_1, \ldots, \alpha_k)$ to the equivalent description $F(\alpha)$ for some $\alpha \in \mathbb{C}$ a lot of information about the structure of the extension may be lost. This will be quite clear in the next section. On the other hand, from a computational point of view the representation of a field as $F(\alpha)$ for an appropriate $\alpha$ with minimal polynomial $p$ is more convenient, since in this case arithmetic in $F(\alpha)$ reduces to arithmetic with polynomials in $F[X]/(p)$.

Let us mention one more useful result that can easily be proven using the distinct embeddings of a field.

**Lemma 2.7** Let $\alpha, \beta$ be algebraic over a field $F$ and let $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$, and $\beta_0, \beta_1, \ldots, \beta_{m-1}$ denote the conjugates of $\alpha$ and $\beta$ over $F$, respectively. Then the conjugates of $\alpha + \beta$ and $\alpha \beta$ are among the complex numbers $\alpha_i + \beta_j$, $\alpha_i \beta_j$, $i = 0, \ldots, n - 1$, $j = 0, \ldots, m - 1$, respectively.

**Proof:** Consider the extension $F(\alpha, \beta)$. Any embedding $\sigma$ of $F(\alpha, \beta)$ over $F$ is uniquely defined by the values of $\sigma(\alpha)$ and $\sigma(\beta)$. Now $\sigma(\alpha) = \alpha_i$ for some conjugate $\alpha_i$ of $\alpha$ and $\sigma(\beta) = \beta_j$ for some conjugate $\beta_j$ of $\beta$.

As has been observed above the conjugates of $\alpha + \beta$ and $\alpha \beta$ are the numbers $\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta)$ and $\sigma(\alpha \beta) = \sigma(\alpha) \sigma(\beta)$, respectively, where $\sigma$ is any embedding of $F(\alpha, \beta)$ over $F$. The lemma follows.

Finally we need from abstract algebra the notions of the trace and norm function of an algebraic extension. Let $E \subset \mathbb{C}$ be an algebraic extension of $F \subset \mathbb{C}$. We consider $E$ as a finite-dimensional vector space over $F$ and fix some basis $\beta_0, \beta_1, \ldots, \beta_{n-1}$ for this vector space. Then for any $\beta \in E$ the mapping

$$
\mu_\beta : E \rightarrow E
$$

$$
\gamma \mapsto \beta \gamma
$$

is an $F$-linear mapping. It therefore has a matrix representation $(c_{ij})$, $i, j = 0, 1, \ldots, n - 1$, with respect to the basis $\{\beta_0, \beta_1, \ldots, \beta_{n-1}\}$, where the $c_{ij}$'s
are elements of $F$ and satisfy
\[ \beta \beta_i = \sum_{j=0}^{n-1} c_{ij} \beta_j, \ i = 0, 1, \ldots, n - 1. \]

The trace $\text{tr}(\beta)$ is the trace of the matrix $(c_{ij})$, that is, the sum of the main diagonal elements, and the norm $\text{no}(\beta)$ is the determinant of this matrix. In particular, the trace and norm map an element of $E$ onto an element of $F$. Of course we may have chosen any basis $B$ for the $F$-vector space $E$ in order to define the trace and norm. It is standard linear algebra that these definitions are invariant under a change of basis.

Next we give a useful alternative definition for the trace and norm function. Consider the distinct field embeddings of $E$ over $F$ which are denoted by $\sigma_i$. Then the trace $\text{tr}_{E:F}$ of $E$ over $F$ can be defined as follows
\[
\text{tr}_{E:F} : E \to F \\
\beta \mapsto \sum_{i=0}^{n-1} \sigma_i(\beta).
\]

Likewise, the norm is defined as
\[
\text{no}_{E:F} : E \to F \\
\beta \mapsto \prod_{i=0}^{n-1} \sigma_i(\beta).
\]

A proof that the definitions are equivalent can be found in [Ja].

From the second definitions it is clear that the trace function is additive and that the norm is multiplicative.

The trace and norm of an element $\beta$ of $E$ are closely related to the minimal polynomial $p_X(\beta)$ of $\beta$ over $F$. If
\[ p_{\beta}(X) = \sum_{i=0}^{m} g_i X^i, \ g_i \in F, \ g_m = 1, \]
is the minimal polynomial of $\beta$ then
\[
(1) \quad \text{tr}(\beta) = (-1)^{n} \frac{n}{m} g_{m-1}
\]
and
\[
(2) \quad \text{no}(\beta) = (-1)^{m} g_0^m.
\]
In particular, if the trace of an element is non-zero then the coefficient $g_{m-1}$ is non-zero, too.

Now let us turn to algebraic number theory. The main objects studied are the algebraic number fields, that is, algebraic extensions $\mathbb{Q}(\alpha)$ of the rationals generated by a single algebraic number over $\mathbb{Q}$. By the Primitive Element Theorem the algebraic number fields are exactly the fields generated by a finite number of algebraic numbers. As mentioned above, if $p(X) = \sum_{i=0}^{n} p_i X^i$, $p_i \in \mathbb{Q}$, $p_n = 1$, is the minimal polynomial of $\alpha$ over $\mathbb{Q}$ then $\mathbb{Q}(\alpha) = \{ \sum_{i=0}^{n-1} q_i \alpha^i | q_i \in \mathbb{Q} \}$ and, moreover, $\mathbb{Q}(\alpha)$ is isomorphic to $\mathbb{Q}[X]/(p)$.

Next to algebraic numbers the most important concepts in algebraic number theory are the concepts of algebraic integers and of the ring of integers of a number field. An algebraic number $\alpha$ is called an algebraic integer or simply an integer if a polynomial $p(X) = \sum p_i X^i$, $p_i \in \mathbb{Z}$, exists such that $p(\alpha) = 0$ and $p_n = 1$. It can be shown using Gauss’ Lemma (see [vdW]) that the minimal polynomial of an algebraic integer is also a monic integer polynomial.

To give some examples, the rational integers $\mathbb{Z}$ are algebraic integers and they are the only algebraic integers contained in the rational numbers. Furthermore $\sqrt[d]{z}$ for arbitrary rational integers $d, z$ is an algebraic integer since it is a root of $p(X) = X^d - z$. Here it does not matter which $d$-th root is meant by $\sqrt[d]{z}$. More generally, and most important for our purposes

**Lemma 2.8** If $\alpha$ is an algebraic integer so is $\sqrt[d]{\alpha}$ for any positive rational integer $d$ and for any of the $d$ possible interpretations of the root.

In fact, if $p(X)$ is the monic minimal polynomial of $\alpha$ then $\sqrt[d]{\alpha}$ is a root of $p(X^d)$, which is also monic.

The set of algebraic integers forms a ring with respect to the usual addition and multiplication in $\mathbb{C}$. This shows that if $\alpha_1, \ldots, \alpha_k$ are algebraic integers then any integer combination $\sum_{i=1}^{k} z_i \alpha_i$, $z_i \in \mathbb{Z}$, of these numbers is also an algebraic integer. Furthermore, to any algebraic number field $\mathbb{Q}(\alpha)$ the set of all algebraic integers contained in $\mathbb{Q}(\alpha)$ can be associated. This set is denoted by $R_\alpha$. Since $R_\alpha$ is the intersection of the field $\mathbb{Q}(\alpha)$ with the ring of all algebraic integers in the complex numbers $R_\alpha$ is also a ring. We give some examples:

1. If $\alpha$ is rational then $\mathbb{Q}(\alpha) = \mathbb{Q}$. As mentioned above the ring of algebraic integers in $\mathbb{Q}$ is $\mathbb{Z}$.
2. Let \( d \in \mathbb{Z} \) be a square-free integer. Then \( \mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} | a, b \in \mathbb{Q}\} \). The ring of integers in \( \mathbb{Q}(\sqrt{d}) \) is given by (see [M])

\[
R_{\sqrt{d}} = \{a + b\sqrt{d} | a, b \in \mathbb{Z}\} \quad \text{if } d \equiv 2, 3 \mod 4
\]

\[
R_{\sqrt{d}} = \left\{a + b \left(1 + \frac{\sqrt{d}}{2}\right) \mid a, b \in \mathbb{Z}\right\} \quad \text{if } d \equiv 1 \mod 4.
\]

The algebraic structure of the ring of integers of an algebraic number field is well-known:

\( R_{\alpha} \) is a free \( \mathbb{Z} \)-module of degree \( n \), the degree of \( \alpha \), i.e., there exist \( n \) algebraic integers \( \beta_i \in R_{\alpha}, i = 0, \ldots, n-1 \), such that any \( \gamma \in R_{\alpha} \) can uniquely be written as \( \gamma = \sum_{i=0}^{n-1} z_i \beta_i, z_i \in \mathbb{Z} \).

Unfortunately, there are no efficient algorithms known (polynomial in the degree of the extension) to compute a basis \( \{\beta_0, \ldots, \beta_{n-1}\} \). On the other hand, for the purpose of this thesis we only need a sub- and a superset of \( R_{\alpha} \), which we are now going to describe. First a useful observation.

Let \( \alpha \) be algebraic and consider the number field \( \mathbb{Q}(\alpha) \) generated by \( \alpha \). We show that there exists an algebraic integer in \( \mathbb{Q}(\alpha) \) that generates the same field and that can easily be computed. Given \( p(X) \), \( \deg p = n \), the minimal polynomial of \( \alpha \), let \( q \) be the least common multiple of the denominators of the coefficients of \( p(X) \). Then \( qa \) is an algebraic integer, since its minimal polynomial is \( q^n p \left(\frac{X}{q}\right) \in \mathbb{Z}[X] \). Furthermore \( \mathbb{Q}(\alpha) = \mathbb{Q}(qa) \). Therefore it can be assumed that algebraic number fields are generated by algebraic integers.

If \( \mathbb{Q}(\alpha) \) is generated by an algebraic integer \( \alpha \) then \( R_{\alpha} \) contains \( \alpha \) and \( \mathbb{Z} \). Since \( R_{\alpha} \) is a ring it follows that \( \mathbb{Z}[\alpha] = \left\{\sum_{i=0}^{n-1} z_i \alpha^i \mid z_i \in \mathbb{Z}\right\} \subseteq R_{\alpha} \). In general, the inclusion is strict. For example, \( \mathbb{Z} \left[\sqrt{d}\right] \) is strictly contained in \( R_{\sqrt{d}} \) if \( d \equiv 1 \mod 4 \). To describe the superset for \( R_{\alpha} \) one more definition is needed.

Let \( \alpha \) be algebraic with minimal polynomial \( p(X) \), \( \deg p = n \). Since \( p(X) \) is irreducible it has \( n \) distinct roots. Denote the conjugates of \( \alpha \) by \( \alpha_0 = \alpha, \alpha_1, \ldots, \alpha_{n-1} \).

**Definition 2.9** The number \( \Delta = \Delta(\alpha) = \prod_{0 \leq i < j \leq n-1} (\alpha_i - \alpha_j)^2 \) is called the discriminant of \( \alpha \) or the discriminant of \( p \).

It is well-known that for an algebraic integer \( \alpha \) its discriminant \( \Delta(\alpha) \) is a rational integer.
A proof of the following lemma which describes the superset of $R_\alpha$ can be found in the book of Marcus [M].

**Lemma 2.10** Let $\alpha$ be an algebraic integer, $Q(\alpha)$ the algebraic number field generated by $\alpha$ and $\Delta$ the discriminant of $\alpha$. Then the ring of integers $R_\alpha$ of $Q(\alpha)$ is contained in the free $\mathbb{Z}$-module $\frac{1}{\Delta}\mathbb{Z} \oplus \frac{1}{\Delta}\mathbb{Z}\alpha \oplus \ldots \oplus \frac{1}{\Delta}\mathbb{Z}\alpha^{n-1}$, i.e., any integer $\gamma$ can uniquely be written as $\gamma = \frac{1}{\Delta}\sum_{i=0}^{n-1} c_i \alpha^i$, $c_i \in \mathbb{Z}$.

In many respects the ring of integers of an algebraic number field has the same properties as the rational integers $\mathbb{Z}$. But there is at least one fundamental difference between $\mathbb{Z}$ and the ring of integers of an algebraic number field. In general, $R_\alpha$ is not a unique factorization domain (UFD). To give an example, consider the field $\mathbb{Q}(\sqrt{-5})$. By Equation (4) 6 can be factored in $R_\alpha$ either as $6 = 2 \cdot 3$ or as $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$. One checks that $2, 3, 1 + \sqrt{-5},$ and $1 - \sqrt{-5}$ are all prime in the ring of integers of $\mathbb{Q}(\sqrt{-5})$.

It is basically due to this fact that the algorithms in the first part of this thesis get rather complicated (and probabilistic) when generalized from $\mathbb{Q}$ to algebraic number fields.

This finishes our description of some fundamental facts in algebra and number theory. In the next section special algebraic extensions called *radical extensions* will be examined in more detail.
3 The Structure of Radical Extensions

In this section we study algebraic extensions of a special type, the so-called radical extensions. Based on the ideas of C. L. Siegel [Si] we give a simplified proof of a theorem that determines the structure of certain radical extensions, for example, those radical extensions contained in the field of real numbers. In particular, it is shown that in this case no non-trivial sum of linearly independent real radicals is itself a radical. Siegel deduced this result from his structure theorem, we proof it directly and base the proof of the structure theorem on this lemma. We also show how to prove and generalize certain results from Kummer Theory using this lemma.

Definition 3.1 Let $F$ be a subfield of the complex numbers $C$. An element $\gamma \in C$ is called a radical over $F$ iff

$$\gamma^d \in F.$$ 

for some positive integer $d$.

Hence radicals are solutions of equations of the form $X^d - \rho, \rho \in F,$ and are therefore algebraic over $F$.

Throughout this thesis we will denote radicals by the familiar symbols $\sqrt[d]{\rho}$. However, the reader should always bear in mind that the symbol $\sqrt[d]{\rho}$ does not uniquely specify a number. This does not cause too many problems in this thesis since the only restriction we frequently impose on a radical is that it should be a real number. In that case we simply say ”the real radical $\sqrt[d]{\rho}$”. Moreover, in this situation we implicitly assume that $\rho$ itself is a real number and that $X^d - \rho = 0$ has a real solution. If $d$ is even we need not worry which of the two possible real solutions to $X^d - \rho = 0$ is meant by the real radical $\sqrt[d]{\rho}$. As will be clear the results will be correct for both of them.

Definition 3.2 An algebraic extension $E$ of $F$ is called a radical extension iff it has the form $E = F(\sqrt[d_1]{\rho_1}, \ldots, \sqrt[d_k]{\rho_k})$ for a finite number of radicals $\sqrt[d_i]{\rho_i}$ over $F$. For $k = 1$ we call the extension a simple radical extension. If $F(\sqrt[d_1]{\rho_1}, \ldots, \sqrt[d_k]{\rho_k}) \subset R$ then the extension will be called a real radical extension.

Recall that if $X^d - \rho \in F[X]$ then the roots of the equation $X^d - \rho = 0$ are given by $\zeta_d \sqrt[d]{\rho}, i = 0, 1, \ldots, d - 1$, where $\zeta_d$ is a primitive $d$-th root of unity and $\sqrt[d]{\rho}$ is any solution of the equation.
The next theorem describes the minimal polynomial of a radical if the field $F$ is a subfield of the reals or if the field $F$ contains “appropriate” roots of unity.

**Theorem 3.3** Let $E \subset \mathbb{C}$ be a field, and let $\sqrt[d]{\rho}$ be a radical over $E$. Assume that either $E \subseteq \mathbb{R}$ and $\sqrt[d]{\rho} \in \mathbb{R}$ or that $E$ contains a primitive $d$-th root of unity. Furthermore assume that $d'$ is the smallest nonnegative integer such that $\sqrt[d]{\rho}^{d'} \in E$. Then $d'$ is a divisor of $d$ and the minimal polynomial $g(X)$ of $\sqrt[d]{\rho}$ over $E$ is given by $g(X) = X^{d'} - \sqrt[d]{\rho}^{d'}$. Hence $d'$ is the degree of $\sqrt[d]{\rho}$ over $E$.

**Proof:** A proof for this theorem can be found in [Mo] or [Si]. For the reader’s convenience we include a proof. We first show that $d'$ divides $d$.

Suppose $d'$ does not divide $d$. In this case we write $d = ld' + k$ with $0 < k < d'$. Since $\sqrt[d]{\rho}^d = \rho \in E$ and $\sqrt[d]{\rho}^{ld'} \in E$ we conclude $\sqrt[d]{\rho}^k \in E$ which contradicts the minimality of $d'$.

Next suppose that $X^{d'} - \sqrt[d]{\rho}^{d'}$ is not the minimal polynomial. Instead assume that the minimal polynomial is given by $f(X) = \sum_{i=0}^{n} f_i X^i$, $f_i \in E$, and $n < d'$. Since the minimal polynomial of an algebraic number $\alpha$ over a field $E$ divides any other polynomial in $E[\alpha]$ with root $\alpha$, $f(X)$ divides $X^d - \rho$. Therefore the constant term $f_0$ of $f$ is a product of $n$ of the roots of $X^d - \rho$. As mentioned, these roots have the form $\zeta^i \sqrt[d]{\rho}$, $i = 0, \ldots, d - 1$, for some primitive $d$-th root of unity $\zeta$. Therefore $f_0$ is of the form $f_0 = \zeta' \sqrt[d]{\rho}^n$ for a $d$-th root of unity $\zeta'$.

If $E \subseteq \mathbb{R}$, $\sqrt[d]{\rho} \in \mathbb{R}$ then $\sqrt[d]{\rho}^n \in \mathbb{R}$. Therefore $f_0/\sqrt[d]{\rho}^n = \zeta' \in \mathbb{R}$. Since $+1$ and $-1$ are the only real roots of unity this implies $\sqrt[d]{\rho}^n \in E$ which contradicts the minimality of $d'$.

Likewise, if $E$ contains a primitive $d$-th root of unity then it contains all $d$-th roots of unity, i.e. $\zeta' \in E$, so $\sqrt[d]{\rho}^n \in E$, again contradicting the minimality of $d'$.

In the analysis of the algorithms that check the linear dependence of radicals the following corollary will be one of the key steps.

**Corollary 3.4** Let $E \subset \mathbb{C}$ be a field, and let $\sqrt[d_1]{\rho_1}$, $\sqrt[d_2]{\rho_2}$ be radicals over $E$. Assume that either $E \subseteq \mathbb{R}$ and $\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2} \in \mathbb{R}$ or that $E$ contains primitive $d_1$-th, $d_2$-th roots of unity. Denote the greatest common divisor (gcd) of $d_1, d_2$ by $d$.

If $\sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \in E$ then $\sqrt[d_1]{\rho_1}^d$, $\sqrt[d_2]{\rho_2}^d \in E$, too.
Proof: In both cases $\sqrt[d]{\rho_1} / \sqrt[d]{\rho_2} \in E$ implies ($\sqrt[d]{\rho_1} / \sqrt[d]{\rho_2}$) \in E. Hence

$$\sqrt[d]{\rho_1}^d = \gamma \sqrt[d]{\rho_2}^d$$

for some $\gamma \in E$.

If $d_1' = d_1 / d$ and $d_2' = d_2 / d$ then Theorem 3.3 shows that the degree of $\sqrt[d]{\rho_1}^{d_1'}$ over $E$ is a divisor of $d_1'$, and that the degree of $\sqrt[d]{\rho_2}^{d_2'}$ is a divisor of $d_2'$. Since $\gcd(d_1', d_2') = 1$ the equality $\sqrt[d]{\rho_1}^{d_1'} = \gamma \sqrt[d]{\rho_2}^{d_2'}$ implies that these degrees must be 1, so $\sqrt[d]{\rho_1}^{d_1'}, \sqrt[d]{\rho_2}^{d_2'} \in E$. 

Our next goal is to describe the form of radicals contained in a radical extension and to strengthen Theorem 3.3 in case the field $E$ is a radical extension of some field $F$ such that $\sqrt[n]{\rho}$ is not only a radical over $E$ but already over $F$.

If $E$ is a radical extension $F(\sqrt[n]{\rho_1}, \ldots, \sqrt[n]{\rho_k})$ of a field $F$ and we denote the degree of the extension $F(\sqrt[n]{\rho_1}, \ldots, \sqrt[n]{\rho_i}) : F(\sqrt[n]{\rho_1}, \ldots, \sqrt[n]{\rho_i}, \sqrt[n]{\rho_{i+1}}, \ldots, \sqrt[n]{\rho_k})$ by $n_i$ then using the results mentioned in the previous section $n = [E : F] = \prod_{i=1}^{k} n_i$ and the field extension has a basis $B$ of the form

$$B = \left\{ \prod_{i=1}^{k} \sqrt[n]{\rho_i}^{e_i} : 0 \leq e_1 < n_1, 0 \leq e_2 < n_2, \ldots, 0 \leq e_k < n_k \right\}.$$ 

Throughout this thesis we refer to this basis as the standard basis of the extension $F(\sqrt[n]{\rho_1}, \ldots, \sqrt[n]{\rho_k}) : F$.

Instead of generating $E$ by the radicals $\sqrt[n]{\rho_1}, \ldots, \sqrt[n]{\rho_k}$ we can also generate this extension by adjoining the radicals $\sqrt[n]{\rho_1}^{-1}, \ldots, \sqrt[n]{\rho_k}^{-1}$ to $F$. Since $F(\sqrt[n]{\rho_1}^{-1}, \ldots, \sqrt[n]{\rho_i}^{-1}, \ldots, \sqrt[n]{\rho_k}^{-1}) = F(\sqrt[n]{\rho_1}, \ldots, \sqrt[n]{\rho_i}, \sqrt[n]{\rho_{i+1}}^{-1}, \ldots, \sqrt[n]{\rho_k}^{-1})$ the degree of $\sqrt[n]{\rho_i}$ over $F(\sqrt[n]{\rho_1}^{-1}, \ldots, \sqrt[n]{\rho_i}^{-1}, \ldots, \sqrt[n]{\rho_k}^{-1})$ is also $n_i$ and hence

$$B' = \left\{ \prod_{i=1}^{k} \sqrt[n]{\rho_i}^{-e_i} : 0 \leq e_1 < n_1, 0 \leq e_2 < n_2, \ldots, 0 \leq e_k < n_k \right\}$$

is another field basis for $E : F$.

Furthermore we need the following lemma.

**Lemma 3.5** Let $E$ be an algebraic extension of a field $F \subset \mathbb{C}$, $[E : F] = n$. By $\tr_{E : F}$ denote the trace function of this extension. If $\{\beta_0, \beta_1, \ldots, \beta_{n-1}\}$ is a $F$-basis for $E$ and $\beta$ a non-zero element of $E$ then

$$\tr_{E : F}(\beta_0 \beta) \neq 0$$
for some \( i \in \{0, 1, \ldots, n-1\} \).

**Proof:** Assume \( \text{tr}_{E:F}(\beta_i \beta) = 0 \) for all \( i \). Since the set \( \{\beta_0 \beta, \beta_1 \beta, \ldots, \beta_{n-1} \beta\} \) is a basis of the extension this implies that the trace function is identically zero. But \( \text{tr}_{E:F}(1) = n \neq 0 \), which proves the lemma. \( \square \)

Despite its simple nature this lemma will prove to be very useful not only in the proof of Siegel’s Theorem but also for the problem of denesting nested radicals.

We are in a position to prove the basic lemma for real radicals.

**Lemma 3.6** Let \( F \) be a real field and let \( \sqrt[n_i]{\rho_i}, i = 1, \ldots, k \), be real radicals over \( F \). Assume \( n_i \) is defined as above. If \( \sqrt[n_k]{\rho_k} \in F(\sqrt[n_1]{\rho_1}, \sqrt[n_2]{\rho_2}, \ldots, \sqrt[n_{k-1}]{\rho_{k-1}}) \) then it has the form

\[
\sqrt[n_k]{\rho_k} = \gamma \prod_{i=1}^{k-1} \sqrt[n_i]{\rho_i^{e_i}}
\]

for some \( \gamma \in F \) and positive integers \( e_i \) satisfying \( 0 \leq e_i < n_i \).

**Proof:** Consider the basis

\[
B' = \left\{ \prod_{i=1}^{k-1} \sqrt[n_i]{\rho_i^{-e_i}, 0 \leq e_1 < n_1, 0 \leq e_2 < n_2, \ldots, 0 \leq e_{k-1} < n_{k-1}} \right\}
\]

of the extension \( E = F(\sqrt[n_1]{\rho_1}, \sqrt[n_2]{\rho_2}, \ldots, \sqrt[n_{k-1}]{\rho_{k-1}}) \) over \( F \). All elements of this basis are clearly radicals over \( F \). By Lemma 3.5 for some element \( \beta \) of this basis the trace \( \text{tr}_{E:F}(\sqrt[n_k]{\rho_k} \beta) \) is non-zero.

If \( g(X) = \sum_{i=0}^{m} g_i X^i \) is the minimal polynomial of \( \sqrt[n_k]{\rho_k} \beta \) over \( F \) then \( g_{m-1} \neq 0 \) (see Equation (2) of Section 2). On the other hand, \( \sqrt[n_k]{\rho_k} \beta \) is a real radical over \( F \). Hence by Theorem 3.3 its minimal polynomial has the form \( g(X) = X^m - (\sqrt[n_k]{\rho_k} \beta)^m \), where \( (\sqrt[n_k]{\rho_k} \beta)^m \in F \). Therefore \( g_{m-1} \) can be non-zero if and only if \( m = 1 \) and \( \sqrt[n_k]{\rho_k} \beta \in F \). This proves the lemma. \( \square \)

To generalize this lemma to complex radicals we need the following well-known result.

**Lemma 3.7** Let \( F \subset \mathbb{C} \) be a field containing primitive \( d_i \)-th roots of unity, \( i = 1, 2, \ldots, k \). Let \( d \) be the least common multiple (lcm) of \( d_i, i = 1, 2, \ldots, k \). Then \( F \) contains a primitive \( d \)-th root of unity.
Proof: We prove the lemma only for $k = 2$ the general case follows by induction on $k$.

Let $\zeta$ be a primitive $d_1d_2$-th root of unity. By Euclid’s algorithm the gcd of $d_1, d_2$ can be written as

$$\gcd(d_1, d_2) = md_1 + nd_2.$$

$\zeta^{\gcd(d_1, d_2)}$ is a primitive $d$-th root of unity. But

$$\zeta^{\gcd(d_1, d_2)} = \zeta^{md_1 + nd_2} = \zeta^{md_1} \zeta^{nd_2}.$$

Since $\zeta^{d_1}$ is a primitive $d_2$-th root of unity and, accordingly, $\zeta^{d_2}$ is a primitive $d_1$-th root of unity, $\zeta^{\gcd(d_1, d_2)}$ is in $F$.

Lemma 3.8 Let $F \subset \mathbb{C}$ be a field and $\sqrt[n_i]{\rho_i}$, $i = 1, \ldots, k$, be radicals over $F$. Furthermore assume that the field $F$ contains primitive $d_i$-th roots of unity. If $\sqrt[n_k]{\rho_k} \in F(\sqrt[n_1]{\rho_1}, \sqrt[n_2]{\rho_2}, \ldots, \sqrt[n_{k-1}]{\rho_{k-1}})$ then $\sqrt[n_k]{\rho_k}$ has the form

$$\sqrt[n_k]{\rho_k} = \gamma \prod_{i=1}^{k-1} d_i^{e_i}$$

for some $\gamma \in F$ and positive integers $e_i$ satisfying $0 \leq e_i < n_i$, where $n_i$ is defined as above.

Proof: The proof is exactly as for Lemma 3.6. Observe that in order to apply Theorem 3.3 and argue that the minimal polynomial of $\sqrt[n_k]{\rho_k} / \beta$, $\beta \in B'$, has the form $X^m - (\sqrt[n_k]{\rho_k} / \beta)^m$ we need a positive integer $d$ such that $(\sqrt[n_k]{\rho_k} / \beta)^d \in F$ and $F$ contains a primitive $d$-th root of unity. The least common multiple of all $d_i$, $i = 1, 2, \ldots, k$, clearly has the first property. By the previous lemma it also has the second property.

Both lemmata can be interpreted as saying that the only radicals in a radical extension of a field $F$ that is either real or contains all relevant roots of unity are the obvious ones, i.e., no non-trivial linear combination of radicals is itself a radical. Throughout the rest of the thesis we will restrict ourselves to radical extensions of fields $F$ that satisfy the conditions of Lemma 3.6 or of Lemma 3.8 and we refer to them as the real and complex case, respectively.
Next we prove two important corollaries to these lemmata. One is a strengthened form of Theorem 3.3. The other corollary describes under which conditions radicals are linearly dependent over a field $F$.

**Theorem 3.9 (Siegel)** Let $F$ be a field and let $d_1\sqrt[\nu]{\rho_1}, d_2\sqrt[\nu]{\rho_2}, \ldots, d_k\sqrt[\nu]{\rho_k}$ be radicals over $F$. In both the real and complex case, as defined above, the minimal polynomial $p_i(X)$ of $\sqrt[\nu]{\rho_i}$ over $F(d_1\sqrt[\nu]{\rho_1}, d_2\sqrt[\nu]{\rho_2}, \ldots, d_{i-1}\sqrt[\nu]{\rho_{i-1}})$ has the form

$$p_i(X) = X^{n_i} - \gamma \prod_{j=1}^{i-1} d_j^{\nu e_j},$$

for some $\gamma \in F$ and integers $e_j \in \mathbb{N}$.

In other words, the degree of the field extension $F(\sqrt[\nu]{\rho_1}, \sqrt[\nu]{\rho_2}, \ldots, \sqrt[\nu]{\rho_i}) : F(\sqrt[\nu]{\rho_1}, \sqrt[\nu]{\rho_2}, \ldots, \sqrt[\nu]{\rho_{i-1}})$ is the smallest number $n_i$ such that $\sqrt[\nu]{\rho_i}$ can be written as a product of an element in $F$ and powers of the radicals $\sqrt[\nu]{\rho_1}, \sqrt[\nu]{\rho_2}, \ldots, \sqrt[\nu]{\rho_{i-1}}$.

**Proof:** By Theorem 3.3 it suffices to show that any power $\sqrt[\nu]{\rho_i^e}$ of $\sqrt[\nu]{\rho_i}$ that is an element of $F(\sqrt[\nu]{\rho_1}, \sqrt[\nu]{\rho_2}, \ldots, \sqrt[\nu]{\rho_{i-1}})$ has the form

$$\rho_i^e = \gamma \prod_{j=0}^{i-1} d_j^{\nu e_j}, e_j \in \mathbb{N}, \gamma \in F.$$

But this is clear from Lemma 3.6 or Lemma 3.8 since any power of $\sqrt[\nu]{\rho_i}$ is a radical over $F$. \hfill \Box

We easily deduce the following corollary.

**Corollary 3.10** Let $F \subset \mathbb{C}$ be a field and let $\sqrt[\nu]{\rho_1}, \sqrt[\nu]{\rho_2}, \ldots, \sqrt[\nu]{\rho_k}$ be non-zero radicals over $F$.

In both the real and complex case a relation $\sum_{i=1}^{k} \kappa_i \sqrt[\nu]{\rho_i} = 0$ for $\kappa_i \in F$ not all zero can exist if and only if different radicals $\sqrt[\nu]{\rho_i}, \sqrt[\nu]{\rho_j}$ exist such that

$$\sqrt[\nu]{\rho_i} = \kappa \sqrt[\nu]{\rho_j}$$

for some $\kappa \in F$.

In other words, the radicals $\sqrt[\nu]{\rho_1}, \ldots, \sqrt[\nu]{\rho_k}$ are linearly independent over $F$ if any two of them are linearly independent.
Proof: We claim that if for every pair of radicals \( \sqrt[\rho_i]{d_i}, \sqrt[\rho_j]{d_j} \)
\[ \sqrt[\rho_i]{d_i} / \sqrt[\rho_j]{d_j} \notin F \]
then the set of radicals \( \{ \sqrt[\rho_1]{d_1}, \sqrt[\rho_2]{d_2}, \ldots, \sqrt[\rho_k]{d_k} \} \) can be extended to a field basis of \( F(\sqrt[\rho_1]{d_1}, \sqrt[\rho_2]{d_2}, \ldots, \sqrt[\rho_k]{d_k}) : F \).

To prove this claim observe that although not all radicals \( \sqrt[\rho_i]{d_i} \) need to be an element of the basis \( B = \{ \prod_{j=1}^{k} \sqrt[\rho_j]{d_j}^{e_j}, \ 0 \leq e_1 < n_1, 0 \leq e_2 < n_2, \ldots, 0 \leq e_k < n_k \} \) (some of the field degrees \( n_j = [F(\sqrt[\rho_1]{d_1}, \sqrt[\rho_2]{d_2}, \ldots, \sqrt[\rho_j]{d_j}) : F(\sqrt[\rho_1]{d_1}, \sqrt[\rho_2]{d_2}, \ldots, \sqrt[\rho_{j-1}]{d_{j-1}})] \) may be 1) by Lemma 3.6 or Lemma 3.8 any radical \( \sqrt[\rho_i]{d_i} \) must be a non-zero multiple of a basis element. The condition \( \sqrt[\rho_i]{d_i} / \sqrt[\rho_j]{d_j} \notin F \) implies that any radical is a multiple of a different basis element. This proves the claim and hence the corollary.

Kneser [K] has given weaker conditions under which the theorem above is correct, i.e., for complex radicals it is not always necessary to assume that the base field \( F \) contains all \( d_i \)-th roots of unity. Unfortunately, even in Kneser’s theorem in general the order of the root of unity that has to be contained in \( F \) is exponential in \( k \). Hence from a computational point of view both versions are only of limited use. As will be seen below they are of interest basically if all \( d_i \)’s are equal\(^7\). But then Kneser’s improved version leads to no significant speed-up in the run times of the algorithms. Moreover, Kneser’s condition is often very hard to check making it infeasible from an algorithmic point of view. Therefore for our purposes it is not worth the effort to state or even proof Kneser’s theorem.

We want to use the previous results to describe all subfields of a radical extension. First consider a field \( F \) containing primitive \( d_i \)-th roots of unity, \( i = 1, 2, \ldots, k \). Let \( E = F(\sqrt[\rho_1]{d_1}, \ldots, \sqrt[\rho_k]{d_k}) \) be a radical extension of \( F \). Denote by \( d \) the least common multiple of the integers \( d_i \). \( F \) contains a primitive \( d \)-th root of unity (see Lemma 3.7). Let \( B = \{ \beta_0, \beta_1, \ldots, \beta_{n-1} \} \) be the standard basis of \( E \). Any element \( \beta_j \in B \) satisfies \( \beta_j^d \in F \).

Consider an arbitrary element \( \gamma \) of \( E \). It generates a subfield of \( E \) and we will show that this subfield is a radical extension of \( F \).

Assume \( \gamma = \kappa_1\beta_1 + \kappa_2\beta_2 + \cdots + \kappa_l\beta_l \) such that the \( \beta_i \)’s are different elements of \( B \) and all coefficients \( \kappa_i \) are non-zero elements of \( F \). We claim

---

\(^7\)This case will be applied in the second part of the thesis where denesting algorithms are considered.
that $F(\gamma) = F(\tilde{\beta}_1, \ldots, \tilde{\beta}_t)$.

By Lemma 2.5, p.17, it suffices to show that $\sigma(\gamma) \neq \tau(\gamma)$ for any pair $(\sigma, \tau)$ of different embeddings of $F(\tilde{\beta}_1, \ldots, \tilde{\beta}_t)$ over $F$.

As mentioned above each $\tilde{\beta}_j$ is a root of a polynomial of the form $X^d - \eta_j$, $\eta_j \in F$. The minimal polynomial of $\tilde{\beta}_j$ must divide this polynomial. Therefore any embedding of this field must send $\tilde{\beta}_j$ to $\zeta_j \tilde{\beta}_j$ for a $d$-th root of unity $\zeta_j$. Hence if

$$\sum_{j=1}^{l} \zeta_j \kappa_j \tilde{\beta}_j \neq \sum_{j=1}^{l} \zeta'_j \kappa_j \tilde{\beta}_j,$$

for $d$-th roots of unity $\zeta_j, \zeta'_j$ such that at least one index $j$ with $\zeta_j \neq \zeta'_j$ exists then the claim follows.

If this is not the case then for certain roots of unity a relation

$$\sum_{j=1}^{l} (\zeta_j - \zeta'_j) \kappa_j \tilde{\beta}_j = 0,$$

exists where not all coefficients $(\zeta_j - \zeta'_j) \kappa_j$ are zero. But all these coefficients are elements in $F$. Hence such a relation contradicts the linear independence of the basis elements $\tilde{\beta}_j$. This proves the claim.

By the Primitive Element Theorem (Theorem 2.6) any subfield of $E$ has the form $F(\gamma)$ for an appropriate $\gamma$ and therefore we have shown that all subfields of $E$ are radical extensions.

Now denote by $\tilde{\rho}_i$ the element $\sqrt[d]{\rho_i} \in F$ and by $A$ the group

$$A = \left\{ \prod_{i=1}^{k} \tilde{\rho}_i^{e_i}, e_i \in \mathbb{Z} \right\}.$$

Furthermore let $F^d$ be the set of all $d$-th powers of elements in $F$. Then $A/F^d$, the set of all elements in $A$ modulo elements in $F^d$, is a finite group. We show that the set of subfields of $E$ is in one-to-one correspondence to the set of subgroups of $A/F^d$. Observe that for any element $\beta_j$ of the standard basis $\beta_j^d$ is in $F$ and is in fact an element of $A$. We want to show that elements in $A/F^d$ are in one-to-one correspondence with the elements of the standard basis.

For any element $\rho$ in $A$ denote by $\sqrt[d]{\rho}$ one of its $d$-th roots. Since $F$ contains a primitive $d$-th root of unity $\tilde{\rho}$ and $F(\sqrt[d]{\rho_1}, \ldots, \sqrt[d]{\rho_k})$ contains $\sqrt[d]{\rho}$ which is a $d$-th root of $\rho_1'$, the field $F(\sqrt[d]{\rho_1}, \ldots, \sqrt[d]{\rho_k})$ will also contain $\sqrt[d]{\rho}$ no matter which $d$-th root of $\rho$ this symbol denotes. Therefore $\sqrt[d]{\rho}$ is
a radical contained in \( F(\sqrt[d_1]{\rho_1}, \ldots, \sqrt[d_k]{\rho_k}) \) and hence must be a product of an element in \( F \) and an element of the standard basis. Finally, observe that for different basis elements \( \beta_i, \beta_j \) the powers \( \beta_i^{d_i}, \beta_j^{d_j} \) generate different equivalence classes in \( A/F^d \) and that two elements \( \rho \) and \( \rho' \) in \( A \) generate the same equivalence class in \( A/F^d \) if and only if the corresponding roots \( \sqrt[d_i]{\rho}, \sqrt[d_j]{\rho'} \) are multiples of the same basis element. This proves the one-to-one correspondence between elements in \( A/F^d \) and elements in the standard basis.

By the proof above any subfield of \( F(\sqrt[d_1]{\rho_1}, \ldots, \sqrt[d_k]{\rho_k}) \) is generated by a subset of the standard basis. But if two subsets generate the same subgroup of \( A/F^d \) then any element in the first group can be written as a product of an element in \( F \) and an element in the second subset, and vice versa. Hence if two subsets generate the same group they generate the same subfield. This proves the one-to-one correspondence between subfields of \( F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}) \) and subgroups of \( A/F^d \).

Assume next that the radicals \( \sqrt[d_i]{\rho_i} \) are linearly independent. It can be shown in exactly the same way as above that any sum \( S = \sum_{i=1}^{k} \kappa_i \sqrt[d_i]{\rho_i}, \kappa_i \in F \setminus \{0\} \), generates the extension \( E = F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}) \). In other words, \( S \) is a primitive element of \( E \). We summarize these observations in

**Theorem 3.11** Let \( F \) be a field containing \( d_i \)-th roots of unity, \( i = 1, 2, \ldots, k \). If \( \sqrt[d_i]{\rho_i}, i = 1, 2, \ldots, k, \) are radicals then all subfields of the radical extension \( E = F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}) \) are radical extensions of \( F \). The subfields are in one-to-one correspondence to subgroups of \( A/F^d \), where \( d = \text{lcm}(d_1, d_2, \ldots, d_k) \), \( A \) is the multiplicative group generated by the elements \( \rho_i^{d_i/d_i} \), and \( F^d \) is the multiplicative group of \( d \)-th powers of elements in \( F \).

If the radicals \( \sqrt[d_i]{\rho_i} \) are linearly independent over \( F \) then any sum \( \sum_{i=1}^{k} \kappa_i \sqrt[d_i]{\rho_i} \) with non-zero coefficients \( \kappa_i \in F \) is a primitive element for \( E \).

The first part of this theorem is of course well-known from Kummer Theory (see [Ar]), where it is proven, however, using Galois theory.

Next we want to prove the same result for real radicals over a real field \( F \). To our knowledge this was not known before. The proof given above cannot be generalized in a straightforward manner to this case since radicals that are linearly independent over a real field \( F \) need not remain linearly independent if we adjoin certain roots of unity to \( F \). The following lemma is the key step to circumvent this problem.

31
Lemma 3.12 Let $F$ be a real field and $\zeta$ a root of unity. If $\sqrt[d]{\rho}$ is a real radical over $F$ contained in $F(\zeta)$ then $\sqrt[d]{\rho}$ is a square root of an element in $F$.

Proof: Assume $\zeta$ is an $m$-th primitive root of unity. The extension $F(\zeta) : F$ is a Galois extension. By Theorem 2.4, p. 16, and Lemma 2.3, p. 16, the Galois group of this extension is isomorphic to a subgroup of $\mathbb{Z}_m^*$. In fact, apply Theorem 2.4 with $K = \mathbb{Q}$, $E = \mathbb{Q}(\zeta)$, and $F = F$. Hence $EF = F(\zeta)$.

Since $\mathbb{Z}_m^*$ is abelian all subfields of $F(\zeta)$ over $F$ are Galois extensions of $F$ (Theorem 2.2). In particular, the extension generated by $\sqrt[d]{\rho}$ must be a Galois extension.

By Theorem 3.3 the minimal polynomial of $\sqrt[d]{\rho}$ over $F$ has the form $X^{d'} - \sqrt[d']{\rho}$ for some $d' \in \mathbb{N}$. If $d' > 2$ then $\sqrt[d]{\rho}$ cannot generate a Galois extension. If it did then $F(\sqrt[d]{\rho})$ had to contain all conjugates of $\sqrt[d]{\rho}$. But for $d' > 2$ some of the conjugates are not even real.

We are now in a position to prove the generalization of Theorem 3.11 to real radical extensions.

Theorem 3.13 Let $F$ be a real field. If $\sqrt[d_i]{\rho_i}, i = 1, 2, \ldots, k$, are real radicals then all subfields of the radical extension $E = F(\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k})$ are radical extensions of $F$. The subfields of $E$ are in one-to-one correspondence to the subgroups of the finite group $A/F^d$, where $d = \text{lcm}(d_1, d_2, \ldots, d_k)$, $A$ is the multiplicative group generated by the elements $\rho_i^{d/d_i}$, and $F^d$ is the multiplicative group of $d$-th powers of elements in $F$.

If the radicals $\sqrt[d_i]{\rho_i}$ are linearly independent over $F$ then any sum $\sum_{i=1}^k \kappa_i \sqrt[d_i]{\rho_i}$ with non-zero coefficients $\kappa_i \in F$ is a primitive element for $E$.

Proof: We claim that if $F$ is a real field and $\sqrt[d_i]{\rho_i}, i = 1, \ldots, k$, are linearly independent real radicals over $F$ then any sum $\sum_{i=1}^k \kappa_i \sqrt[d_i]{\rho_i}$, $\kappa_i \in F$, $\kappa_i \neq 0$, generates the extension $F(\sqrt[d_i]{\rho_i}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k})$.

Applying this claim to the standard basis of a real radical extension it implies as in the complex case that any subfield of a real radical extension is itself a radical extension. The one-to-one correspondence stated in the theorem can then be proven in exactly the same way as in the complex case by identifying an element of $A$ with one of its real $d$-th roots.

To prove the claim denote by $d$ the least common multiple of the integers $d_i$. Let $\zeta_d$ be a $d$-th primitive root of unity.
By the previous lemma if 

\[ \frac{\kappa_j \sqrt[p_j]{q}}{\kappa_i \sqrt[p_i]{q}} \in F(\zeta_d) \]

for two different indices \( i, j \leq k \) then the ratio must be a square root \( \sqrt{\mu} \) of some element \( \mu \) in \( F \). Hence after an appropriate renumbering of the radicals \( \sqrt[p_i]{q} \) the sum \( \sum_{i=1}^{k} \kappa_i \sqrt[p_i]{q} \) can be written as

\[ \sum_{i=1}^{l} \kappa_i \left( 1 + \sqrt{\mu_{i,1}} + \cdots + \sqrt{\mu_{i,j_i}} \right) \sqrt[p_i]{q}, \]

where each \( \sqrt[p_{i,h}]{q} \in F(\zeta_d) \) is a square root of an element in \( F \) and the radicals \( \sqrt[p_i]{q}, i = 1, 2, \ldots, l \), are linearly independent over \( F(\zeta_d) \) (Corollary 3.10).

The elements in a set \( \{1, \sqrt[p_{i,1}]{q}, \ldots, \sqrt[p_{i,j_i}]{q}\}, i = 1, 2, \ldots, l \), are linearly independent over \( F \). If this were not the case then by Corollary 3.10 the ratio of two elements in this set would be an element of \( F \). But this would imply

\[ \frac{\kappa_j \sqrt[p_j]{q}}{\kappa_i \sqrt[p_i]{q}} \in F, i \neq j, i, j \leq k, \]

for at least one ratio of radicals. This contradicts the linear independence of these radicals.

In particular, the sums \( 1 + \sqrt[p_{i,1}]{q} + \cdots + \sqrt[p_{i,j_i}]{q} \) are non-zero for all \( i = 1, 2, \ldots, l \).

Next observe that \( E = F(\sqrt[p_1]{q}, \sqrt[p_2]{q}, \ldots, \sqrt[p_k]{q}) \) is the same field as the field generated by the elements in

\[ G = \bigcup_{i=1}^{l} \{ \sqrt[p_i]{q}, \sqrt[p_{i,1}]{q}, \ldots, \sqrt[p_{i,j_i}]{q} \}. \]

Any embedding of \( E \) maps \( \sqrt[p_i]{q} \) onto \( \zeta_i \sqrt[p_i]{q} \) for some \( d_i \)-th root of unity \( \zeta_i \) and \( \sqrt[p_{i,h}]{q} \) is mapped either onto \( \sqrt[p_{i,h}]{q} \) or onto \( -\sqrt[p_{i,h}]{q} \). Furthermore different embeddings map at least one element in \( G \) onto different complex numbers.

Hence by Lemma 2.5, p.17, it suffices to show that

\[ \sum_{i=1}^{l} \zeta_i \kappa_i \left( 1 + \epsilon_{i,1} \sqrt[p_{i,1}]{q} + \cdots + \epsilon_{i,j_i} \sqrt[p_{i,j_i}]{q} \right) \sqrt[p_i]{q} \neq \]

33
\[ \sum_{i=1}^{l} \zeta_i' \kappa_i \left( 1 + \epsilon_{i,1} \mu_{i,1} + \cdots + \epsilon_{i,j_i} \mu_{i,j_i} \right) \sqrt[p_i]{\rho_i}, \]

where \( \zeta_i, \zeta_i' \) are \( d \)-th roots of unity, \( \epsilon_{i,h}, \epsilon_{i,h}' \in \{ +1, -1 \} \), and for at least one index \( i \) \( \zeta_i \neq \zeta_i' \) or \( \epsilon_{i,h} \neq \epsilon_{i,h}' \) for some \( h \).

Observe that in both sums the coefficients are elements of \( F(\zeta_d) \). Therefore if the two sums are equal then a linear relation over \( F(\zeta_d) \) between the radicals \( \sqrt[p_i]{\rho_i}, i = 1, \ldots, l \), exists. By construction these radicals are linearly independent over \( F(\zeta_d) \) and hence the two sums above are equal if and only if all coefficients of the difference of the sums are zero. We will show that this is impossible.

Let \( i \) be such that \( \zeta_i \neq \zeta_i' \) or \( \epsilon_{i,h} \neq \epsilon_{i,h}' \) for some \( h \). \( \kappa_i \neq 0 \) by assumption.

Hence

\[ \zeta_i \left( 1 + \epsilon_{i,1} \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i} \sqrt[p_i]{\mu_{i,j_i}} \right) - \zeta_i' \left( 1 + \epsilon_{i,1}' \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i}' \sqrt[p_i]{\mu_{i,j_i}} \right) \]

must be zero.

If \( \zeta_i = \zeta_i' \) then

\[ 1 + \epsilon_{i,1} \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i} \sqrt[p_i]{\mu_{i,j_i}} = 1 + \epsilon_{i,1}' \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i}' \sqrt[p_i]{\mu_{i,j_i}}. \]

But for a fixed \( i \) the \( \sqrt[p_i]{\mu_{i,h}} \)'s are linearly independent over \( F \) and for at least one index the coefficients \( \epsilon_{i,h}, \epsilon_{i,h}' \in \{ +1, -1 \} \) are different. Therefore the two sums cannot be equal.

Again because for a fixed \( i \) the \( \sqrt[p_i]{\mu_{i,h}} \) and 1 are linearly independent \( 1 + \epsilon_{i,1} \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i} \sqrt[p_i]{\mu_{i,j_i}} \) is non-zero for any combination of signs \( \epsilon_{i,h} \). Hence if \( \zeta_i \neq \zeta_i' \) then

\[ \frac{\zeta_i}{\zeta_i'} = \frac{1 + \epsilon_{i,1} \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i} \sqrt[p_i]{\mu_{i,j_i}}}{1 + \epsilon_{i,1}' \sqrt[p_i]{\mu_{i,1}} + \cdots + \epsilon_{i,j_i}' \sqrt[p_i]{\mu_{i,j_i}}}. \]

But the expression on the right-hand side is real. Since \( \zeta_i' \) is a \( d \)-th root of unity this implies \( \zeta_i' = -\zeta_i \).

So in this case the coefficient can be zero if and only if

\[ 2 + (\epsilon_{i,1} + \epsilon_{i,1}') \sqrt[p_i]{\mu_{i,1}} + \cdots + (\epsilon_{i,j_i} + \epsilon_{i,j_i}') \sqrt[p_i]{\mu_{i,j_i}} = 0. \]

This is again impossible since the elements in \( \{ 1, \sqrt[p_i]{\mu_{i,1}}, \ldots, \sqrt[p_i]{\mu_{i,j_i}} \} \) are linearly independent over \( F \) and the coefficient in front of 1 is non-zero. This finally proves the claim and hence the theorem. \( \blacksquare \)
Theorem 3.11 and Theorem 3.13 show how to construct a primitive element for a radical extension if the generators of the extension are linearly independent. This result will be used in Section 7 where we analyze our denesting algorithms.

Corollary 3.10, on the other hand, provides us with a simple way to check the linear dependence of a set of radicals \{\sqrt[n_1]{\rho_1}, \ldots, \sqrt[n_k]{\rho_k}\} over a field satisfying one of the conditions in Corollary 3.10. In fact, it has only to be tested whether any ratio \( \sqrt[n_i]{\rho_i} / \sqrt[n_j]{\rho_j} \) of radicals is contained in \( F \). In the next sections we show how to solve this problem efficiently. To demonstrate the basic ideas we first restrict ourselves to the rather simple case \( F = \mathbb{Q} \). 

35
4 Radicals over the Rational Numbers

In the first part of this section we show how to decide efficiently whether a set of real radicals \( R = \{ \sqrt[2]{q_1}, \ldots, \sqrt[k]{q_k} \} \) over \( \mathbb{Q} \) is linearly independent. By efficient we mean an algorithm that is polynomial in the input size of the set \( R \).

If \( d_i \) is even and \( q_i \) is positive then the symbol \( \sqrt[q_i]{d_i} \) does not specify uniquely a real number. To avoid ambiguity and simplify the notation in this case we assume that \( \sqrt[q_i]{d_i} \) is positive. Note that this is no restriction as far as we are interested in linear combinations of radicals because we may change the sign of the coefficients accordingly. With this convention the input size of \( \sqrt[q_i]{d_i} \) is the input size of \( d_i, q_i \). If the sum of the bit size of the numerator and denominator in \( q_i \) and of the bit size of \( d_i \) is at most \( l \) we call \( \sqrt[q_i]{d_i} \) an \( l \)-bit radical. So the input size of a set \( R \) containing \( k \) \( l \)-bit radicals is bounded by \( O(kl) \) and an algorithm that decides in time polynomial in the input size whether a set of real radicals is linearly independent has to be polynomial in the number of radicals \( k \) and in the logarithm of the degrees \( d_i \). Due to Corollary 3.4 and Corollary 3.10 we will achieve such a run time by a rather simple algorithm.

If we want to check whether a linear combination of the radicals above with coefficients in \( \mathbb{Q} \) is zero we transform the sum into a linear combination of radicals that are linearly independent. The sum will be zero if and only if the coefficients in the transformed sum are zero. The transformation is done by combining those radicals whose ratio is rational. Due to the solution to the first problem this transformation is easily computed.

In the second part of this section we apply the results of the first part to the problem of determining the sign of sums of square roots. Although these results cannot be used to describe an algorithm that computes the sign in polynomial time, in many cases they can be used to speed up the algorithms known so far.

4.1 Linear Dependence of Radicals over the Rational Numbers

Given a set of real radicals \( \{ \sqrt[q_1]{d_1}, \sqrt[q_2]{d_2}, \ldots, \sqrt[q_k]{d_k} \} \) over the rational numbers we want to check whether these radicals are linearly independent. Likewise, we want to determine whether a given rational combination of these radicals with coefficients in \( \mathbb{Q} \) is zero.

Due to Corollary 3.10 we only have to check whether for any pair of
radicals $\sqrt[n]{q_i}$, $\sqrt[n]{q_j}$, their ratio is a rational number. As will be seen later, by Corollary 3.4 this property in turn can be tested by determining for three radicals over $\mathbb{Q}$ whether they are rational. Moreover, the degree of these radicals will be smaller than $\max\{d_i, d_j\}$. We can restrict the problem even further to roots of integers.

**Lemma 4.1.1** Let $q \in \mathbb{Q}$, $q = \frac{a}{b}$, $\gcd(a, b) = 1$ and $d \in \mathbb{N}$. Then $\sqrt[n]{q} \in \mathbb{Q}$ if and only if $\sqrt[n]{a} \in \mathbb{Z}$ and $\sqrt[n]{b} \in \mathbb{Z}$.

**Proof:** Assume $\sqrt[n]{q} = \frac{\sqrt[n]{a}}{\sqrt[n]{b}} \in \mathbb{Q}$. Hence $\sqrt[n]{ab^{d-1}} \in \mathbb{Q}$. From the unique factorization property of $\mathbb{Z}$ it easily follows that if $\sqrt[n]{z} \in \mathbb{Q}$ for $z \in \mathbb{Z}$ then $\sqrt[n]{z} \in \mathbb{Z}$. Therefore $\sqrt[n]{ab^{d-1}} \in \mathbb{Z}$. Since $\gcd(a, b) = 1$ this implies $\sqrt[n]{a} \in \mathbb{Z}$ and $\sqrt[n]{b^{d-1}} \in \mathbb{Z}$.

Due to the unique factorization $\sqrt[n]{b^{d-1}} \in \mathbb{Z}$ if and only if every prime dividing $b^{d-1}$ does so with exponent divisible by $d$. Any such exponent has the form $e(d - 1)$, where $e$ is the exponent with which the prime divides $b$. But $\gcd(d - 1, d) = 1$, so if $d$ divides $e(d - 1)$ it must already divide $e$. Therefore $\sqrt[n]{b^{d-1}} \in \mathbb{Z}$ implies $\sqrt[n]{b} \in \mathbb{Z}$, too. 

We now show how to check in polynomial time whether an integer is a $d$-th power.

**Lemma 4.1.2** Let $z, d$ be $l$-bit integers. It can be decided with $\mathcal{O}(l)$ elementary operations on integers of bit size at most $\mathcal{O}(l)$ whether $\sqrt[n]{z} \in \mathbb{Z}$ and, if so, within the same time bounds it can be computed.

**Proof:** We may assume $d \leq \log z$. Otherwise $X^d - z = 0$ has a solution in $\mathbb{Z}$ if and only if $z = 1$ or $z = -1$. Furthermore we can restrict ourselves to positive integers.

The integer $z' \in \mathbb{Z}$ such that if $\sqrt[n]{z} \in \mathbb{Z}$ then $\sqrt[n]{z} = z'$ can easily determined by a binary search on the interval $I = [0, 2^{\lceil \frac{l}{d} \rceil}]$. In each step of the binary search we have to determine whether an element $\hat{z}$ from $I$ if raised to the $d$-th power is smaller or larger than $z$ or equal to $z$. The $d$-th power is computed by successive squaring. Also observe that we may stop when a power of $\hat{z}$ has been computed that is larger than $z$. Hence the binary search can be done using at most $\mathcal{O}(\frac{l}{d} \log d) \in \mathcal{O}(l)$ elementary operations on integers of size $\mathcal{O}(l)$. 

37
The binary search described above can be considered as an approximation algorithm for $\sqrt{z}$. When we generalize the results of this section to number fields we will see that by applying a more sophisticated approximation algorithm the run time in Lemma 4.1.2 can be reduced to $O(\log l)$ elementary operations on floating-point numbers of size $O(l)$.

We combine the previous lemmata to deduce the following theorem.

**Theorem 4.1.3** Let $d_1\sqrt{q_1}$, $d_2\sqrt{q_2}$ be l-bit radicals over $\mathbb{Q}$. It can be decided using $O(l)$ elementary operations on integers of length $O(l)$ whether the ratio of these radicals is in $\mathbb{Q}$. Furthermore if the ratio is rational it can be computed within the same time bounds.

**Proof:** First the gcd of $d_1$ and $d_2$ is computed. By Schönhage’s result [Sc1] this can be done by $O(\log l)$ operations on $O(l)$-bit integers. Denote the gcd by $d$.

We also compute within these time bounds $d_i' := d_i/d$.

By Corollary 3.4, p.24, if $d_1\sqrt{q_1}/d_2\sqrt{q_2} \in \mathbb{Q}$ then $d_i' = d_1' = q_i' \in \mathbb{Q}$, $i = 1, 2$. We check whether this is the case by applying the algorithm leading to Lemma 4.1.2 to $d_i'$ and the numerator and denominator of $q_i$, $i = 1, 2$. The correctness of this procedure follows from Lemma 4.1.1, and by Lemma 4.1.2 it uses $O(l)$ elementary operations on integers of size $O(l)$.

Then we compute $q_1'/q_2'$ and determine (using again Lemma 4.1.1, Lemma 4.1.2) whether $d_1\sqrt{q_1'/q_2'} \in \mathbb{Q}$. Since $d_1\sqrt{q_1}/d_2\sqrt{q_2} = d_1\sqrt{q_1'/q_2'}$, this will give the desired result. As the numerators and denominators of $q_1'$, $q_2'$ have at most $l$ bits the theorem follows.

Combining this result with Corollary 3.10 leads to

**Corollary 4.1.4** Let $\{d_1\sqrt{q_1}, d_2\sqrt{q_2}, \ldots, d_k\sqrt{q_k}\}$ be a set of l-bit radicals over $\mathbb{Q}$. It can be decided using $O(k^2l)$ elementary operations on integers of length $O(l)$ whether this set is linearly independent over $\mathbb{Q}$. Within the same time bound a maximal subset of linearly independent radicals can be computed.

**Proof:** Apply the algorithm of the previous theorem to the $k(k-1)/2$ different ratios of radicals.
We can also use Theorem 4.1.3 above to check whether a linear combination of radicals \( S = \sum_{i=1}^{k} v_i \sqrt[\alpha]{q_i} \) is zero.

**Corollary 4.1.5** Let \( \sqrt[\alpha]{q_i}, i = 1, \ldots, k \) be \( l \)-bit radicals. If \( v_i \in \mathbb{Q}, i = 1, \ldots, k, \) are such that the numerator and denominator of \( v_i \) are \( l \)-bit integers then it can be decided using \( O(k^2l) \) elementary operations on integers of length \( O(l) \) and \( O(k) \) elementary operations on integers of length \( O(kl) \) whether the sum \( S = \sum_{i=1}^{k} v_i \sqrt[\alpha]{q_i} \) is zero.

**Proof:** Using the algorithm of the previous corollary partition \( R = \{ \sqrt[\alpha]{q_1}, \sqrt[\alpha]{q_2}, \ldots, \sqrt[\alpha]{q_h} \} \) into subsets \( R_1, \ldots, R_h \) such that two radicals are in the same subset if and only if their ratio is rational. To simplify the notation assume \( \sqrt[\alpha]{q_i} \in R_i, i = 1, \ldots, h. \) This partitioning can be done by \( O(k^2l) \) elementary operations on integers of length \( O(l) \). Within the same time bounds rational numbers \( r_{ij} \) are computed such that if \( \sqrt[\alpha]{q_j} / \sqrt[\alpha]{q_i} \in \mathbb{Q} \) then \( \sqrt[\alpha]{q_j} / \sqrt[\alpha]{q_i} = r_{ij}. \) Hence

\[
S = \sum_{i=1}^{k} v_i \sqrt[\alpha]{q_i} = \sum_{i=1}^{h} \left( \sum_{\sqrt[\alpha]{q_j} \in R_i} v_j r_{ij} \right) \sqrt[\alpha]{q_i}.
\]

Since for any pair of radicals in \( R' = \{ \sqrt[\alpha]{q_1}, \sqrt[\alpha]{q_2}, \ldots, \sqrt[\alpha]{q_h} \} \) their ratio is not a rational number, by Corollary 3.10, p.28, \( S = 0 \) if and only if

\[
\sum_{\sqrt[\alpha]{q_j} \in R_i} v_j r_{ij} = 0, \text{ for } i = 1, \ldots, h.
\]

To compute these sums for each \( i \) between 1 and \( h \) we compute the product of the denominators of the rational numbers in \( \{ v_j r_{ij} | j \text{ such that } \sqrt[\alpha]{q_j} \in R_i \} \). Observe that each denominator has at most \( O(l) \) bits. Hence this can be done by \( O(k) \) elementary operations on integers of size at most \( O(kl) \). Once the product has been computed the sum is easily determined by \( O(k) \) elementary operations on integers of size \( O(kl) \).

Corollary 4.1.5 is restricted to real radicals. We could apply the same idea to, say, a sum of complex fifth roots if we work with the fifth cyclotomic field instead of \( \mathbb{Q} \). In fact, then we could apply the results of Section 3 to complex radicals. But this implies that we have to check efficiently whether a ratio of complex radicals is contained in a cyclotomic field. So this problem...
naturally leads to the question whether the results of this section can be
generalized to algebraic number fields.

Before we answer this question we describe an application of Corollary
4.1.5 to the problem of determining the sign of a sum of square roots.

4.2 Comparing Sums of Square Roots

Probably one of the most important open problems in algebraic computing
is the determination of the sign of a sum $S = \sum_{i=1}^{k} c_i \sqrt{n_i}$, $c_i \in \mathbb{Z}$, $n_i \in \mathbb{N}$, of
square roots in polynomial time. This problem occurs for example in shortest
path computations if the bit complexity is considered. Another application
is the Euclidean Traveling Salesman Problem again if bit arithmetic instead
of infinite-precision arithmetic is used.

It seems that the only algorithm known so far for this problem is the
brute force solution, that is, compute $S$ with precision good enough to de-
termine the sign. Then the question is how many bits of $S$ have to be com-
puted. Unfortunately, even the best bounds are exponential in the number
$k$ of square roots.

To prove such a bound is actually quite simple. In fact, consider the
algebraic integer $S = \sum_{i=1}^{k} c_i \sqrt{n_i}$, $c_i \in \mathbb{Z}$, $c_i \neq 0$. By the previous subsection
we may assume that the square roots $\sqrt{n_i}$ are linearly independent over $\mathbb{Q}$
and that $S$ is non-zero.

The polynomial

$$g(X) = \prod \left( X - \sum_{i=1}^{k} \epsilon_i c_i \sqrt{n_i} \right),$$

where the product is over all $k$-tuples $(\epsilon_1, \epsilon_2, \ldots, \epsilon_k) \in \{+1, -1\}^k$, is an
integer polynomial with non-zero constant coefficient.

First observe that it must be a polynomial with rational coefficients. In
fact, by Lemma 2.7, p. 18, the conjugates of a sum $\sum_{i=1}^{k} \epsilon_i c_i \sqrt{n_i}$ are also of
the form $\sum_{i=1}^{k} \epsilon' c_i \sqrt{n_i}$ with $\epsilon' \in \{-1, +1\}$. Hence if $\alpha$ is a root of $g$, so are
all its conjugates over $\mathbb{Q}$. This implies that $g$ is a polynomial with rational
coefficients.

Furthermore each $\sum_{i=1}^{k} \epsilon_i c_i \sqrt{n_i}$ is an algebraic integer. This implies that
the coefficients of $g$ are algebraic integers. Since the coefficients are rational
they must be rational integers. Finally observe that each sum $\sum_{i=1}^{k} \epsilon_i c_i \sqrt{n_i}$
is non-zero since the $\sqrt{n_i}$’s are linearly independent over $\mathbb{Q}$ and the $c_i$’s,$\epsilon_i$’s
are non-zero. Hence the constant coefficient of $g$ must be non-zero, too.
Assume $|c_i| < 2^l$, $n_i < 2^l$, for all $i$. Then the coefficients in $g$ are bounded in absolute value by $(2^{2l}k)^{2l}$ (see also Lemma 7.1.1, p. 142, to be proven later). Now we may use the following result that goes back to Cauchy (see also [Mi1]).

**Lemma 4.2.1** Let $p(X) = \sum_{i=0}^{n} p_iX^i$, $p_0 \neq 0$, $p_n \neq 0$ be a polynomial with complex coefficients. Then any root $z$ of $p$ satisfies

(i) $|z| < 1 + \max\{|p_0|,|p_1|,...,|p_{n-1}|\}/|p_n|$

(ii) $|z| > |p_0|/|p_0| + \max\{|p_1|,|p_2|,...,|p_n|\}$

Applying (ii) to the polynomial $g$ above shows that if $S = \sum_{i=1}^{k} c_i\sqrt{n_i} \neq 0$ then $|S| \geq (2^{2l}k)^{-2l} - 1$ hence $\mathcal{O}(2^l(l + \log k))$ bits suffice in order to determine the sign of $S$.

But the arguments above show a bit more. In fact, assume that the degree of the minimal polynomial $g_0$ of $\sum_{i=1}^{k} c_i\sqrt{n_i}$ is $N \leq 2^k$. Then (Lemma 2.7, p.18), $g_0$ has the form

$$g_0 = \prod_{(\epsilon_1,...,\epsilon_k) \in H} \left( X - \sum_{i=1}^{k} \epsilon_i c_i\sqrt{n_i} \right),$$

where $H$ is a subset of $\{+1,-1\}^k$ of size $N$. Hence the coefficients of $g_0$ are bounded in absolute value by $(k2^{2l})^N$. This proves that $2N(l + \log k)$ bits suffice to determine the sign of $S = \sum_{i=1}^{k} c_i\sqrt{n_i}$.

Of course $N$ may be $2^k$ and in this case knowing $N$ will lead to no speed-up in the algorithm that determines the sign of $S = \sum_{i=1}^{k} c_i\sqrt{n_i}$. On the other hand, if $N$ is much smaller than $2^k$ computing $N$ in advance may save a lot of time. In the rest of this section we show how to compute the degree $N$ in time that is almost proportional to the degree itself.

Recall that we assume that the square roots $\sqrt{n_i}$ are linearly independent over $\mathbb{Q}$. Hence by Theorem 3.13, p.32, the degree of $S$ is the same as the degree of the extension $\mathbb{Q}(\sqrt{n_1},...\sqrt{n_k})$ over $\mathbb{Q}$.

Applying the theorems of Section 3 and the algorithms of the previous subsection a basis for the extension and hence the degree is easily determined as follows.

Assume w.l.o.g. that no square root is rational. Then a basis for $\mathbb{Q}(\sqrt{n_1})$ over $\mathbb{Q}$ is given by $B_1 = \{1, \sqrt{n_1}\}$. Furthermore a basis for $\mathbb{Q}(\sqrt{n_1}, \sqrt{n_2})$ is $B_2 = \{1, \sqrt{n_1}, \sqrt{n_2}, \sqrt{n_1n_2}\}$. Recall that $\sqrt{n_2} \notin \mathbb{Q}(\sqrt{n_1})$ since $\sqrt{n_1}/\sqrt{n_2} \notin \mathbb{Q}$.
Q and by Lemma 3.6, p.26, \( \sqrt{m_2} \in Q(\sqrt{m_1}) \) would imply either \( \sqrt{m_2} \in Q \) or \( \sqrt{m_2} = q\sqrt{m_1} \), \( q \in Q \).

Similarly Lemma 3.6 implies \( \sqrt{m_3} \in Q(\sqrt{m_1}, \sqrt{m_2}) \) if and only if \( \sqrt{m_3} \) is a rational multiple of an element in \( B_2 = \{1, \sqrt{m_1}, \sqrt{m_2}, \sqrt{m_1m_2}\} \). In which case a basis for \( Q(\sqrt{m_1}, \sqrt{m_2}, \sqrt{m_3}) \) is given by \( B_2 \). Otherwise a basis is given by \( B_3 = B_2 \cup B_2\sqrt{m_3} \) where \( B_2\sqrt{m_3} \) denotes the set we get by multiplying every element in \( B_2 \) by \( \sqrt{m_3} \).

Continuing in this way we see that once we have a basis \( B_{i-1} \) for the extension \( Q(\sqrt{m_1}, \ldots, \sqrt{m_{i-1}}) : Q \) consisting of square roots only a basis \( B_i \) for \( Q(\sqrt{m_1}, \ldots, \sqrt{m_i}) : Q \) is given either by \( B_{i-1} \) or by \( B_{i-1} \cup B_{i-1}\sqrt{m_i} \) depending on whether \( \sqrt{m_i} \) is a rational multiple of an element in \( B_{i-1} \) or not.

Therefore a basis \( B_k \) for \( Q(\sqrt{m_1}, \ldots, \sqrt{m_k}) : Q \) can be computed by checking for at most \( k|B_k| = kN \) pairs of square roots of integers whether their ratio is rational. Furthermore the absolute value of the integers involved is bounded by \( \prod_{i=1}^{k} n_i \). Hence by Theorem 4.1.3 the degree \( N \) of \( Q(\sqrt{m_1}, \ldots, \sqrt{m_k}) : Q \) together with a basis \( B_k \) can be computed using \( O(k^2Nl) \) elementary operations on integers of size \( O(kl) \). Here as above \( n_i < 2^l \).

**Theorem 4.2.2** Let \( \{\sqrt{m_1}, \sqrt{m_2}, \ldots, \sqrt{m_k}\} \), \( n_i \in N, n_i < 2^l \), be a set of real square roots. The degree \( N \) of the extension \( Q(\sqrt{m_1}, \sqrt{m_2}, \ldots, \sqrt{m_k}) : Q \) can be determined using \( O(k^2Nl) \) elementary operations on integers of size \( O(kl) \).

For any sum \( S = \sum_{i=1}^{k} c_i\sqrt{m_i}, c_i \in Z, |c_i| < 2^l \), its degree over \( Q \) can be determined within the same time bound.

**Proof:** By Corollary 4.1.5 within the time bounds stated we compute a maximal subset of linearly independent radicals in \( \{\sqrt{m_1}, \sqrt{m_2}, \ldots, \sqrt{m_k}\} \) and transform the sum \( S \) into a sum of linearly independent radicals.

In order to compute the degree of the extension we apply the process described above to the maximal subset of linearly independent radicals. In order to determine the degree of \( S \) we apply the process to those linearly independent radicals that have a non-zero coefficient in the transformed sum (see Theorem 3.13, p. 32).

Combining this result with the arguments above leads to

**Corollary 4.2.3** For any sum \( S = \sum_{i=1}^{k} c_i\sqrt{m_i}, c_i \in Z, n_i \in N, n_i, |c_i| < 2^l \), its sign can be determined by \( O(k^2Nl) \) elementary operations on integers
of size $O(\mathcal{O}(kl))$ and $O(k)$ elementary operations on floating-point numbers of size $O(N(l + \log k))$. Here $N$ is the degree of $S$ over $\mathbb{Q}$.

**Proof:** By the previous theorem the degree of $S$ over $\mathbb{Q}$ can be determined within the time bounds stated. In particular, it will be determined whether $S$ is zero. If this is not the case, by the remarks above $2N(l + \log k)$ correct bits of $S$ suffice to determine the sign of $S$. As is easily seen computing each square root $\sqrt{n_i}$ with absolute error less than $2^{-3N(l+\log k)}$, multiplying the approximations with the corresponding coefficient $c_i$, and adding the results leads to an approximation to $S$ as required.

Computing each square root $\sqrt{n_i}$ with relative error less than $2^{-4N(l+\log k)}$ on the other hand yields an approximation to $\sqrt{n_i}$ as required. Since taking square roots is one of the elementary operations the corollary follows.

Let us briefly sketch how to generalize the results above to arbitrary sums of radicals $\sum_{i=1}^{k} c_i \sqrt[di]{n_i}$ with $d_i, n_i \in \mathbb{N}$ and $c_i \in \mathbb{Z}$. To prove the next corollary we need certain results from the following sections. The proof of these lemmata needs some terminology that can be defined only in a later section therefore we will only mention in the proof which lemmata are used.

**Corollary 4.2.4** Let $S = \sum_{i=1}^{k} c_i \sqrt[di]{n_i}$, $d_i, n_i \in \mathbb{N}, c_i \in \mathbb{Z}$, be a sum of real radicals. Assume $n_i, d_i, |c_i| < 2^l$. Using $O((k^2 + kN^2 \log N)l)$ elementary operations on integers of size $O((k + N \log N)l)$ and $O(k \log(Nl \log k))$ elementary operations on floating-point numbers of size $O(N(l + \log k))$ the sign of $S$ can be determined. Here $N$ is the degree of $S$ over $\mathbb{Q}$.

**Proof:** By Corollary 4.1.5 within the time bounds stated $S$ can be transformed into a sum of linearly independent radicals. In particular, it can be decided whether $S$ is zero. If this is not the case, we determine the degree of the radical extension generated by the radicals appearing in the transformed sum with non-zero coefficient. By Theorem 3.13, p. 32, the degree of this extension is the degree of $S$. As will follow from the proof of Lemma 7.2.1, p. 135, this degree can be determined within the time bounds stated (see also Remark 7.2.2, p. 140).

As was the case for the square roots the minimal polynomial of $S$ is of the form

$$g_0 = \prod_{(\zeta_1, \ldots, \zeta_k) \in H} \left( X - \sum_{i=1}^{k} \zeta_i c_i \sqrt{\zeta_i} \right),$$

43
where $H$ is a set of $k$-tuples $(\zeta_1, \zeta_2, \ldots, \zeta_k)$ such that $\zeta_i$ is $d_i$-th root of unity. The cardinality of $H$ is $N$.

As before, the absolute value of the coefficients in $g_0$ is bounded by $(k2^{2l})^N$ and by Cauchy’s bound $2N(l + \log k)$ bits suffice to determine the sign of $S = \sum_{i=1}^{k} c_i \sqrt[n]{n_i}$. As in the previous lemma approximating each radical $\sqrt[n]{n_i}$ with absolute error less than $2^{-3N(l+\log k)}$, multiplying the approximations with the corresponding coefficient $c_i$, and adding the results leads to an approximation to $S$ as required. Lemma 5.4.6, p. 71, adjusted to the rational case shows that the approximations to $\sqrt[n]{n_i}$ can computed by $O(k \log(Nl \log k))$ elementary operations on floating-point numbers of size $O(N(l + \log k))$. This proves the corollary. □
5 Linear Dependence of Radicals over Algebraic Number Fields

In this section we generalize the results of Section 4.1 to real radicals over real algebraic number fields and complex radicals over number fields containing certain roots of unity. That is, we want to describe efficient algorithms that check whether a set of radicals \( \{ \sqrt[\rho_1]{d_1}, \sqrt[\rho_2]{d_2}, \ldots, \sqrt[\rho_k]{d_k} \} \) over an algebraic number field \( \mathbb{Q}(\alpha) \) is linearly dependent or whether a given linear combination of these radicals is zero. By Corollary 3.10, p. 28, if \( \alpha \) and the radicals \( \sqrt[\rho_i]{d_i} \) are real or \( \mathbb{Q}(\alpha) \) contains primitive \( d_i \)-th roots of unity this problem can be reduced to the question whether a ratio of radicals \( \sqrt[\rho_1]{d_1}/\sqrt[\rho_2]{d_2} \) is contained in \( \mathbb{Q}(\alpha) \). So it basically remains to describe an algorithm that answers questions of this kind efficiently. Since the ring of integers in \( \mathbb{Q}(\alpha) \) is in general not a unique factorization domain we really have to work with ratios and cannot, like in Section 4, restrict ourselves to roots of algebraic integers.

First a few words on the input of the problems above. We assume that the field \( \mathbb{Q}(\alpha) \) is given by the \( n \)-tuple \( (p_{n-1}, p_{n-2}, \ldots, p_1, p_0) \in \mathbb{Z}^n \) such that the algebraic integer \( \alpha \) has minimal polynomial \( p(X) = X^n + \sum_{i=0}^{n-1} p_i X^i \). We also assume that \( \alpha \) is distinguished from its conjugates by an isolating interval \( I \) in the real case and an isolating rectangle \( R \) in the complex case. That is, an interval or rectangle that contains exactly one root of \( p \). We will see below (see Lemma 5.1.3, p. 49) that the endpoints can be chosen as rational numbers such that both numerator and denominator are \( O(nl) \)-bit integers. This representation of an algebraic number field is exactly the one that has been used by Loos [Lo1].

In algebraic number theory usually an element \( \beta \in \mathbb{Q}(\alpha) \) is described by an \( n \)-tuple \( (q_0, q_1, \ldots, q_{n-2}, q_{n-1}) \in \mathbb{Q}^n \) such that \( \rho = \sum_{i=0}^{n-1} q_i \alpha^i \). In this thesis we assume instead that \( \rho \) is encoded by an \( (n+1) \)-tuple \( (b, b_0, b_1, \ldots, b_{n-2}, b_{n-1}) \in \mathbb{Z}^{n+1} \) such that \( \rho = \frac{1}{b} \sum_{i=0}^{n-1} b_i \alpha^i \). Moreover, we assume \( \gcd(b, b_0, b_1, \ldots, b_{n-1}) = 1 \) and \( b > 0 \). The integer \( b \) will be called the denominator of \( \rho \). By computing the lcm of the denominators of the \( q_i \) we can easily go from one representation to the other. In any case, the total input size of \( \rho \) is a linear in \( n \) and the maximum bit size of the integers \( p_i, b, b_i \).

To avoid ambiguity, in the algorithms of this section we assume that a radical \( \sqrt[\rho]{d} \) is given by \( d, \rho \), and a positive integer \( k \) between 0 and \( d - 1 \)

---

\(^6\)Recall that we can restrict ourselves to fields generated by integers.
such that
\[ \sqrt[d]{\rho} = \zeta_d^k |\rho|^{\frac{1}{d}} \left( \cos \frac{1}{d} \phi + i \sin \frac{1}{d} \phi \right), \]
where \( \phi \in (-\pi, \pi] \) denotes the angle of \( \rho \) when written in polar coordinates, \(|\rho|^{\frac{1}{d}}\) is the positive \( d \)-th root, and \( \zeta_d = \cos \frac{2\pi}{d} + i \sin \frac{2\pi}{d} \). For the sake of brevity, the integer \( k \) will not be mentioned explicitly. If we require that the radical is real we assume that the \( k \) is an appropriate integer, i.e., \( k = 0 \) if \( d \) is odd and \( k = 0 \) or \( k = \frac{d}{2} \) if \( d \) is even.

Therefore the input size of a set of radicals \{ \( \sqrt[d]{\rho_1}, \sqrt[d]{\rho_2}, \ldots, \sqrt[d]{\rho_k} \) \} over an algebraic number field is polynomial in the degree of the field, in the bit size of the coefficients of the minimal polynomial of \( \alpha \), in the bit size of the coefficients of the \( \rho_i \)'s, and in \( \log d_i \), \( i = 1, 2, \ldots, k \).

Remark that if \( \mathbb{Q}(\alpha) \subset \mathbb{R} \) then the question whether two elements \( \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2} \in \{ \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k} \} \subset \mathbb{R} \) have a ratio contained in \( \mathbb{Q}(\alpha) \), and hence the question whether the elements in \{ \( \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k} \) \} are linearly independent, does not depend on the values of the radicals in the set as long as they are real. If \( \mathbb{Q}(\alpha) \) contains primitive \( d_i \)-th roots of unity then the question whether the elements in \{ \( \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k} \) \} are linearly independent is independent of the values of the radicals. We need not even distinguish \( \alpha \) from its conjugates by an isolating rectangle since all conjugate fields contain \( d_i \)-th roots of unity. For linear combinations on the other hand, we really need to specify the roots \( \sqrt[d]{\rho} \) and the element \( \alpha \) since for some interpretations of the roots a sum of radicals may be zero and for others not.

In a sense the question whether the set of radicals is linearly independent over a field containing \( d_i \)-th roots of unity is the only purely algebraic problem. It can be solved purely symbolically, that is, we need to specify \( \alpha \) only by its minimal polynomial and the radicals \( \sqrt[d]{\rho} \) by \( \rho_i \) and \( d_i \). We will apply a test whether \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \in \mathbb{Q}(\alpha) \) for complex radicals \( \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2} \) only for fields \( \mathbb{Q}(\alpha) \) containing primitive \( d_i \)-th roots of unity. Hence \( \mathbb{Q}(\alpha) \) has degree at least \( \max\{\varphi(d_1), \varphi(d_2)\} \). Recall that \( \varphi \) denotes Euler’s \( \varphi \)-function. Since \( \varphi(d) = \Omega\left(\frac{d}{\log d} \right) \) (see [Ap]) and the algorithm that decides whether a ratio of radicals is in \( \mathbb{Q}(\alpha) \) must be polynomial in the degree of the extension we cannot hope for an algorithm that is polynomial in \( \log d_i \), instead we have to be content with an algorithm that is polynomial in \( d_i \). But then the question whether \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \in \mathbb{Q}(\alpha) \) can be solved by factoring \( \rho_2^{d_2} X^{d_1} - \rho_1^{d_1} \) over \( \mathbb{Q}(\alpha) \) using a polynomial time factorization algorithm for polynomials over algebraic number fields (see [La],[Le]).
For real radicals over real algebraic number fields the situation is different. We will describe an algorithm with run time polynomial in $\log d_1, \log d_2$ that tests for a ratio of radicals $\sqrt[\nu_1]{\rho_1}/\sqrt[\nu_2]{\rho_2}, \rho_1, \rho_2 \in \mathbb{Q}(\alpha)$, whether it is contained in $\mathbb{Q}(\alpha)$ and, if so, computes the representation in $\mathbb{Q}(\alpha)$.

So we have to derive an upper bound on the representation size of $\sqrt[\nu_1]{\rho_1}\sqrt[\nu_2]{\rho_2}$ as an element of $\mathbb{Q}(\alpha)$ that is polynomial in $\log d_i$. In the second subsection of this section we derive a bound that is even independent of $d_1$ and $d_2$. This may seem quite surprising but to some extent it only reflects the well-known fact that the size of the coefficients of a factor of a polynomial depends only on the degree of the factor and on the coefficients of the original polynomial but not on its degree (see [Mi1]).

The algorithm for ratios of real radicals which we describe has a similar structure as the algorithm for real radicals over $\mathbb{Q}$. In the first step an approximation to the ratio $\sqrt[\nu_1]{\rho_1}/\sqrt[\nu_2]{\rho_2}$ will be computed that is good enough to determine in a second step an element $\gamma \in \mathbb{Q}(\alpha)$ such that if $\sqrt[\nu_1]{\rho_1}\sqrt[\nu_2]{\rho_2}$ is in $\mathbb{Q}(\alpha)$ then it must be $\gamma$. Finally, in the third step we check whether $\gamma$ equals the ratio of radicals by computing $\gamma^{d_1d_2}$ and comparing it to $\rho_1^{d_2}/\rho_2^{d_1}$.

We will describe the second step first in order to determine the quality of the approximation to $\sqrt[\nu_1]{\rho_1}/\sqrt[\nu_2]{\rho_2}$ that allows us to determine a unique element $\gamma \in \mathbb{Q}(\alpha)$ such that if $\sqrt[\nu_1]{\rho_1}/\sqrt[\nu_2]{\rho_2} \in \mathbb{Q}(\alpha)$ then it must be $\gamma$. The main ingredient to this step will be a variant of the algorithm of Kannan et al. [KLL] to determine the minimal polynomial of an algebraic number once an approximation to this number is given. That these problems are closely related is not hard to see because both can be described as searching for a linear dependence between algebraic numbers. Reconstructing the minimal polynomial of an algebraic number is the same as constructing a minimal integer linear dependence between its powers. Likewise, for (re-)discovering the exact representation of an element $\gamma \in \mathbb{Q}(\alpha)$ we have to determine an integer linear dependence between $\gamma$ and powers of $\alpha$.

Next we describe how to approximate $\sqrt[\nu_1]{\rho_1}/\sqrt[\nu_2]{\rho_2}$ with precision as required by the second step. This approximation algorithm will be based on algorithms due to R. Brent who showed how to evaluate exp, ln, and the trigonometric functions efficiently with small relative error.

In the third step probabilistic arguments are used to check whether the number $\gamma$ computed by the first two steps satisfies $\rho_2^{d_1}\gamma^{d_1d_2} = \rho_1^{d_2}$. If $\gamma^{d_1d_2}$ and $\rho_1^{d_2}, \rho_2^{d_1}$ are computed by successive squaring then it may happen that the coefficients in the results have $\Omega(d_1d_2)$ bits. Moreover, unlike the rational
case the size of $d_1, d_2$ cannot be polynomially bounded from above by the representation size of $\rho_1, \rho_2$. In our approach $\gamma^{d_1d_2}, \rho_1^{d_2}$, and $\rho_2^{d_1}$ are actually computed be successive squaring, after each multiplication step, however, we reduce the coefficients of the resulting elements in $\mathbb{Q}(\alpha)$ by several randomly chosen small integers. Finally we compare whether the elements in $\mathbb{Q}(\alpha)$ thus obtained are equal. It will be shown that with high probability this gives the correct result.

By not using the reduction step we get a deterministic test with runtime polynomial in $d_1, d_2$. We can apply this algorithm in the complex case where it yields a more efficient solution to the question whether a ratio of radicals is in $\mathbb{Q}(\alpha)$ than using a factorization algorithm.

As a special case, when $d_2 = 1$, the algorithms can be used to determine efficiently whether a radical $\sqrt[\rho]{\rho}, \rho \in \mathbb{Q}(\alpha)$, is in $\mathbb{Q}(\alpha)$.

In the last subsection we apply the previous results to sums of radicals.

### 5.1 Definitions and Bounds

We review some of the basic facts about polynomials. For the reader’s convenience we include some proofs.

Let $p = \sum_{i=0}^{n} p_i X^i \in \mathbb{C}[X]$ be a polynomial with complex coefficients. The length $|p|_2$ of $p$ denotes the euclidean length $(\sum_{i=0}^{n} |p_i|^2)^{\frac{1}{2}}$ of the vector $(p_0, \ldots, p_n)$. The height $|p|_\infty$ of $p$ is the $L_\infty$-norm $\max\{|p_0|, |p_1|, \ldots, |p_n|\}$ of $(p_0, \ldots, p_n)$. Observe that for any irreducible integer polynomial $p$ its length is at least $\sqrt{2}$.

The first result we mention is a generalization of Cauchy’s bounds (Lemma 4.2.1, p.41) which is due to Landau (see Mignotte [Mi1] for a proof).

**Definition 5.1.1** If $p(X) = \sum_{i=0}^{n} p_i X^i \in \mathbb{C}[X]$ has roots $\alpha_0, \ldots, \alpha_{n-1}$, then the measure $M(p)$ of $p$ is defined as

$$M(p) = |p_n| \prod_{j=0}^{n-1} \max\{1, |\alpha_j|\}$$

**Lemma 5.1.2 (Landau)** The measure $M(p)$ of a polynomial $p$ satisfies

$$M(p) \leq |p|_2.$$
Next we give the well-known root separation bound for polynomials. It determines for example how to choose the isolating intervals for the description of the fields $\mathbb{Q}(\alpha)$.

**Lemma 5.1.3** If $p$ is an integer polynomial of degree $n$ then any pair $\alpha_1, \alpha_2$ of distinct roots of $p$ satisfies

$$|\alpha_1 - \alpha_2| > n^{-(n+2)/2} |p|^{1-n}.$$

A proof for this fact can be found in Mignotte’s overview [Mi1].

For later purposes we also derive an upper bound for the absolute value of the discriminant. We get this via Hadamard’s bound for the determinant of a matrix. For a polynomial with distinct roots $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$, we defined the discriminant $\Delta$ as

$$\Delta = \prod_{0 \leq i < j \leq n-1} (\alpha_i - \alpha_j)^2.$$

Hence $|\Delta|^{\frac{1}{2}} = |\prod_{0 \leq i < j \leq n-1} (\alpha_i - \alpha_j)|$ is the absolute value of the determinant of

$$\begin{bmatrix}
1 & \alpha_0 & \ldots & \alpha_0^{n-1} \\
1 & \alpha_1 & \ldots & \alpha_1^{n-1} \\
\vdots & & & \\
1 & \alpha_{n-1} & \ldots & \alpha_{n-1}^{n-1}
\end{bmatrix},$$

which is a Vandermonde matrix.

**Lemma 5.1.4 (Hadamard’s bound)** Let $M = (m_{ij})$ be $(n \times n)$-square matrix whose entries are complex numbers. Define $H(M)$ as the product of the euclidean norm of the rows of $M$,

$$H(M) = \prod_{i=1}^{n} \left( \sum_{j=1}^{n} |m_{ij}|^2 \right)^{\frac{1}{2}}.$$

Then

$$|\det M| \leq H(M).$$

Combining Landau’s bound for the measure of a polynomial with Hadamard’s bound yields an upper bound for the discriminant of a polynomial $p$. 

49
Lemma 5.1.5 Let \( p(X) = \sum_{i=0}^{n} p_{i} X^{i} \) be an irreducible polynomial with complex coefficients. Then the discriminant \( \Delta \) of \( p \) satisfies

\[
|\Delta|^{\frac{1}{2}} < |p_{n}|^{-(n-1)n^{2}}|p|_{2}^{n}.
\]

Proof: Let \( \alpha_{0}, \alpha_{1}, \ldots, \alpha_{n-1} \) be the roots of \( p \). \( |\Delta|^{\frac{1}{2}} = |\det D| \), where \( D \) is the matrix

\[
\begin{pmatrix}
1 & \alpha_{0} & \cdots & \alpha_{0}^{n-1} \\
1 & \alpha_{1} & \cdots & \alpha_{1}^{n-1} \\
\vdots & & & \\
1 & \alpha_{n-1} & \cdots & \alpha_{n-1}^{n-1}
\end{pmatrix}
\]

from above.

By Hadamard’s bound

\[
|\det D| \leq \left( \prod_{i=0}^{n-1} \sum_{j=0}^{n-1} |\alpha_{i}|^{2j} \right)^{\frac{1}{2}}.
\]

Expanding the product yields \( n^{2} \) terms each of which is smaller than \( \frac{1}{|p_{n}|^{2(n-1)}}(M(p))^{2(n-1)} \). Hence by Landau’s bound for \( M(p) \)

\[
|\Delta|^{\frac{1}{2}} \leq n^{\frac{3}{2}} \frac{1}{|p_{n}|^{n-1}}|p|_{2}^{n-1}.
\]

Hadamard’s bound can be generalized to matrices with complex polynomials as entries. Below this bound will be applied to the resultant of two polynomials.

Lemma 5.1.6 (Goldstein-Graham) Let \( M(X) = (M_{ij}(X)) \) be an \( n \times n \)-matrix whose entries are polynomials with complex coefficients. Denote by \( m_{ij} \) the \( L_{1} \)-norm of \( M_{ij}(X) \), that is, the sum of the absolute values of the coefficients of \( M_{ij} \). Furthermore define \( M' \) as \( M' = (m_{ij}) \). Then the polynomial \( \det M(X) \in \mathbb{C}[X] \) satisfies

\[
|\det M(X)|_{2} \leq H(M').
\]

Next we deduce some bounds on the representation size of algebraic numbers.
For an algebraic number $\beta = \beta_0$ of degree $m$ its infinity norm $[\beta]_\infty$ is defined as $\max\{|\beta_0|, |\beta_1|, \ldots, |\beta_{m-1}|\}$, the maximum of the absolute values of its conjugates.

If we assume that $\beta$ is an element of the algebraic number field $\mathbb{Q}(\alpha)$, $\beta = \frac{1}{b} \sum_{i=0}^{n-1} b_i \alpha^i$, $\gcd(b, b_0, b_1, \ldots, b_{n-1}) = 1$, and denote the distinct field embeddings of $\mathbb{Q}(\alpha)$ over $\mathbb{Q}$ by $\sigma_j$, $j = 0, \ldots, n - 1$, then $[\beta]_\infty = \max\{|\sigma_0(\beta)|, |\sigma_1(\beta)|, \ldots, |\sigma_{n-1}(\beta)|\}$.

Let us note the following important property of the infinity norm. It will be used throughout the thesis.

**Lemma 5.1.7** Let $\alpha, \beta$ be algebraic numbers. Then

(i) $[\alpha + \beta]_\infty \leq [\alpha]_\infty + [\beta]_\infty$,

(ii) $[\alpha\beta]_\infty \leq [\alpha]_\infty [\beta]_\infty$.

**Proof:** Denote the conjugates of $\alpha$ and $\beta$ by $\alpha_i, i = 0, 1, \ldots, n - 1$ and $\beta_j, j = 0, 1, \ldots, m - 1$, respectively. By Lemma 2.7, p.18, the conjugates of $\alpha + \beta$ and $\alpha\beta$ are among the numbers $\alpha_i + \beta_j$ and $\alpha_i\beta_j$, respectively. Hence

$$[\alpha + \beta]_\infty \leq \max\{|\alpha_i + \beta_j|\} \leq \max\{|\alpha_i|\} + \max\{|\beta_j|\} = [\alpha]_\infty + [\beta]_\infty,$$

and similarly for the product. \qed

For $\beta \in \mathbb{Q}(\alpha)$ as above $[\beta]$ is defined as $\max\{|b|, |b_0|, |b_1|, \ldots, |b_{n-1}|\}$. We will call $[\beta]$ the coefficient size of $\beta$ with respect to $\mathbb{Q}(\alpha)$. Observe that $\beta$ has in general not a unique representation as $\beta = \frac{1}{b} \sum_{i=0}^{n-1} b_i \alpha^i$, $b, b_i \in \mathbb{Z}$, therefore we cannot define $[\beta]$ without any restrictions on the integers $b, b_i$.

By our definition for $[\beta]$ we can speak of the coefficient size of an algebraic number rather than of the coefficient size of a representation for an algebraic number.

$[\beta]$ depends on the field $\mathbb{Q}(\alpha)$ and the generator $\alpha$. But in general it will be clear which field $\mathbb{Q}(\alpha)$ and which generator we are referring to, therefore we do not mention them explicitly in the symbol $[\beta]$.

With these definitions we get

**Lemma 5.1.8** Let $\mathbb{Q}(\alpha)$ be an algebraic number field, where $\alpha$ is an algebraic integer with minimal polynomial $p(X) = \sum_{i=0}^{n} p_i X^i$, $p_n = 1$, $p_i \in \mathbb{Z}$. Then any element $\beta$ of the ring of integers $R_\alpha$ of $\mathbb{Q}(\alpha)$ satisfies

$$[\beta] < n^{2n} |p|_2^{2n} [\beta]_\infty.$$
Proof: Since $\beta \in R_\alpha$ it can be represented as $\beta = \frac{1}{\Delta} \sum_{i=0}^{n} b_i \alpha^i$, $b_i \in \mathbb{Z}$ (see Lemma 2.10, p. 22). Hence $\Delta \beta = \sum_{i=0}^{n} b_i \alpha^i$, and the integers $b_i$, $i = 0, \ldots, n-1$, are the unique solution to the following equation

$$
\begin{bmatrix}
1 & \alpha_0 & \ldots & \alpha_0^{n-1} \\
1 & \alpha_1 & \ldots & \alpha_1^{n-1} \\
\vdots & & & \\
1 & \alpha_{n-1} & \ldots & \alpha_{n-1}^{n-1}
\end{bmatrix}
\begin{bmatrix}
b_0 \\
b_1 \\
\vdots \\
b_{n-1}
\end{bmatrix}
= 
\begin{bmatrix}
\Delta \sigma_0(\beta) \\
\Delta \sigma_1(\beta) \\
\vdots \\
\Delta \sigma_{n-1}(\beta)
\end{bmatrix}.
$$

A bound on the absolute values of $\Delta, b_i$ is clearly an upper bound for $[\beta]$.

Denote the matrix on the left-hand side of the equation above by $D$. Since $(\det D)^2$ is the discriminant $\Delta$ of $\alpha$, $\det D \neq 0$. Hence $D^{-1}$ is defined and

$$
\begin{bmatrix}
b_0 \\
b_1 \\
\vdots \\
b_{n-1}
\end{bmatrix}
= D^{-1}
\begin{bmatrix}
\Delta \sigma_0(\beta) \\
\Delta \sigma_1(\beta) \\
\vdots \\
\Delta \sigma_{n-1}(\beta)
\end{bmatrix}.
$$

Let $D^{-1} = (d_{ij})$.

Then $b_i = \Delta \sum_{j=0}^{n-1} d_{ij} \sigma_j(\beta)$ or $|b_i| \leq |\Delta| \sum_{j=0}^{n-1} |d_{ij}| |\sigma_j(\beta)|$. So we need an upper bound on $|d_{ij}|$.

It is standard linear algebra that

$$
d_{ij} = (-1)^{i+j} \det D_{ji}/ \det D,
$$

where $D_{ji}$ is the $(n-1) \times (n-1)$-matrix we get by deleting the $j$-th row and $i$-th column in $D$. As in the proof of Lemma 5.1.5 using Hadamard’s bound and Landau’s estimate for the measure yields

$$
|\det D_{ji}| \leq (n-1)^{n-1} |p|^{n-1}_2,
$$

because $p_n = 1$. Hence

$$
|d_{ij}| < n^{n-1} |p|^{n-1}_2/|\Delta^{1/2}|.
$$

The lemma follows from the previous bound on $|\Delta|^{1/2}$. 

Formulated differently the lemma has already been shown by Weinberger, Rothschild [WR] and others (see [La1], [Le], [LMc]).

Next we recall from Loos’ paper on resultants [Lo1] how to construct for an element $\beta = \frac{1}{\Delta} \sum_{i=0}^{n-1} b_i \alpha^i$, $b_i \in \mathbb{Z}$, an integer polynomial whose roots are the conjugates of $\beta$ and deduce bounds on $[\beta]_\infty$ and $[\beta^{-1}]_\infty$. 

52
Definition 5.1.9 Let \( f(X) = \sum_{i=0}^{n} f_i X^i \) and \( g(X) = \sum_{i=0}^{m} g_i X^i \) be polynomials over an arbitrary commutative ring \( R \). The determinant of the following \((n + m) \times (n + m)\)-matrix \( M \) that consists of \( m \) rows of shifted coefficients of \( f \) and \( n \) rows of shifted coefficients of \( g \)

\[
M = \begin{bmatrix}
\vdots & f_n & f_{n-1} & \cdots & f_1 & f_0 \\
\vdots & f_n & f_{n-1} & \cdots & f_1 & f_0 \\
g_m & g_{m-1} & \cdots & g_1 & g_0 \\
g_m & g_{m-1} & \cdots & g_1 & g_0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
g_m & g_{m-1} & \cdots & g_1 & g_0
\end{bmatrix}
\]

is called the resultant \( \text{res}(f, g) \) of \( f \) and \( g \).

Lemma 5.1.10 (Loos) Let \( \beta = \frac{1}{b} \sum_{i=0}^{n-1} b_i \alpha^i, b, b_i \in \mathbb{Z} \) be an element of \( \mathbb{Q}(\alpha) \). Denote by \( B(Y) \) the polynomial \( B(Y) = \sum_{i=0}^{n-1} b_i Y^i \). Then the resultant \( r(X) \in \mathbb{Z}[X] \) of \( bX - B(Y) \) and \( p(Y) \) taken with respect to the ring \( \mathbb{Z}[X] \) has roots \( \sigma_j(\beta), j = 0, \ldots, n - 1 \). Here \( p \) is the minimal polynomial of \( \alpha \) and the \( \sigma_j \)'s are the distinct field embeddings of \( \mathbb{Q}(\alpha) \).

Using the polynomial \( r \) we can prove an upper bound on \( [\beta^{-1}]_\infty \).

Lemma 5.1.11 Let \( \mathbb{Q}(\alpha) \) and \( \beta \) be as above. Then

(i) \( \beta \) and \( \beta^{-1} \) are roots of polynomials \( r \) and \( r' \), respectively, whose 2-norm is bounded by

\[
|r|_2 = |r'|_2 < n^{2n} |p|_2^n [\beta]^n
\]

(ii)

\[
[\beta]_\infty < 3 |\beta| |p|_2^n
\]

and

\[
[\beta^{-1}]_\infty < n^{2n} |p|_2^n [\beta]^n.
\]
**Proof:** We may choose \( r \) as the polynomial from the previous lemma. Moreover, if
\[
r(X) = \sum_{i=0}^{n} r_i X^i, \quad r_i \in \mathbb{Z},
\]
then \( \beta^{-1} \) is a root of
\[
r'(X) = \sum_{i=0}^{n} r_{n-i} X^i.
\]
Hence \( |r|_2 = |r'|_2 \). To get the bound on \( |r|_2 \) apply the Graham-Goldstein bound to the matrix defining \( r \) which shows
\[
|r|_2 < |p|_2^{\frac{n}{2}} \left( \sum_{i=1}^{n-1} |b_i|^2 + (|b_0| + |b|)^2 \right)^{\frac{n}{2}} \leq |p|_2^{n} (n + 3)^{\frac{n}{2}} |\beta|^n < n^{2n} |p|_2 |\beta|^n.
\]
By Landau’s bound for the measure of a polynomial (Lemma 5.1.2) for any integer polynomial \( p \) its roots are bounded in absolute value by \( |p|_2 \). Applying this to the polynomial \( r' \) and the bound on the length of \( r' \) from above proves the bound for \( |\beta^{-1}|_\infty \).

To prove the better bound on \( |\beta|_\infty \) note that by Landau’s bound on the measure of \( p \) \( |\alpha|_\infty \leq |p|_2 \). Moreover observe that \( p \) is irreducible hence its length is at least \( \sqrt{2} \). \( |\beta|_\infty \leq |\beta| \max \{ \sum_{i=0}^{n-1} |\sigma_j(\alpha)|^i \} \), where the maximum is over all field embeddings \( \sigma_j \) of \( \mathbb{Q}(\alpha) \). Since
\[
\sum_{i=0}^{n-1} |\sigma_j(\alpha)|^i \leq \sum_{i=1}^{n-1} |p|_2^i \leq \frac{|p|_2^n - 1}{|p|_2 - 1},
\]
the bound follows from \( |p|_2 - 1 \geq \sqrt{2} - 1 > \frac{1}{3} \).

Note that Lemma 5.1.8 and Lemma 5.1.11 are dual to one another in the sense that the first one bounds \( |\beta| \) in terms of \( |\beta|_\infty \) and the second one bounds \( |\beta|_\infty \) and \( |\beta^{-1}|_\infty \) in terms of \( |\beta| \). We will use both directions in the sequel.
5.2 Lattice Basis Reduction and Reconstructing Algebraic Numbers

In this section we answer the following questions:

Given an approximation $\gamma$ to an element $\gamma \in \mathbb{Q}(\alpha)$ and a guarantee that the integer coefficients $c, c_i$ in $\gamma = \frac{1}{c} \sum_{i=0}^{n-1} c_i \alpha^i$ are bounded in absolute value by some positive number $K$. Can the coefficients $c, c_i$ be computed exactly? How good do we have to choose the approximation $\gamma$?

We will show that a variant of the Kannan, Lenstra, Lovász algorithm to reconstruct minimal polynomials (see [KLL]) can be used to solve this problem.

The main tool that will be used are lattices. Given a set $V$ of linearly independent vectors $V = \{v_1, \ldots, v_m\} \subset \mathbb{R}^n$, $m \leq n$, the lattice $\Lambda(V)$ generated by these vectors is the set

$$\Lambda(V) = \left\{ \sum_{i=1}^{m} z_i v_i \mid z_i \in \mathbb{Z} \right\}$$

of vectors that can be written as linear integer combinations of the vectors in $V$. The vectors $v_i$ will be described as the columns of an $(n \times m)$-matrix which is also called $V$.

Since we assume that the columns in $V$ are linearly independent any vector $v \in \Lambda(V)$ can be identified with a unique vector $(z_1, \ldots, z_m) \in \mathbb{Z}^m$ such that $v = \sum_{i=1}^{m} z_i v_i$. Using the matrix $V$ this reads as $v = V(z_1, \ldots, z_m)^T$.

We give two examples for lattices.

**Example 1** The identity matrix $I_n$ in $\mathbb{R}^n$

$$I_n = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$

generates the set of all vectors in $\mathbb{R}^n$ with integer coordinates. The same lattice is generated if one column in $I_n$ is replaced by a column consisting only of 1’s.

**Example 2** The $2 \times 2$-matrix

$$\begin{bmatrix} -1 & +1 \\ +1 & +1 \end{bmatrix}$$

generates all vectors $(a, b)^T \in \mathbb{R}^2$ satisfying $a \equiv b \pmod{2}$. 

55
An important concept is a basis of a lattice. A subset of a lattice \( \Lambda(V) \) is called a basis if every element in \( \Lambda(V) \) can uniquely be written as an integer linear combination of the basis vectors. As can be seen from the first example above and is clear from linear algebra a lattice may have many different bases. On the other hand, each lattice has a basis.

Another important concept that has been introduced by Lenstra et al. [LLL] in their break-through work on polynomial factorization is a so-called reduced basis of lattice. For a set \( V \) of linearly independent vectors we call the set of vectors we get by applying the Gram-Schmidt orthogonalization process to \( V \) the Gram-Schmidt-version of \( V \).

**Definition 5.2.1** Let \( \Lambda(V) \subset \mathbb{R}^n \) be an \( m \)-dimensional lattice. A basis \( \{b_1, \ldots, b_m\} \) for \( \Lambda(V) \) is called reduced if its Gram-Schmidt version \( \{b_1^*, \ldots, b_m^*\} \) has the following two properties

(i) \[
\left| (b_i, b_j^*)/(b_j^*, b_j^*) \right| \leq \frac{1}{2}, 1 \leq j < i \leq m,
\]

(ii) \[
\|b_j^*\|^2 \leq 2\|b_{j+1}^*\|^2, j \leq m,
\]

where \((,\)\) is the usual inner product and \(\|\cdot\|_2\) denotes the euclidean length. The basis \( \{b_1, \ldots, b_m\} \) is called semi-reduced if its Gram-Schmidt version satisfies

\[
\|b_i^*\|^2 \leq 2^{m+j-i}\|b_j^*\|^2, 1 \leq i < j \leq m.
\]

The concept of a semi-reduced basis was introduced by Schönhage [Sc4].

Observe that a lattice is a discrete object. So the length of a shortest vector taken with respect to the euclidean length \(\|\cdot\|_2\) is uniquely defined although there may be many different vector of this length. For a reduced basis the next lemma was originally proven in [LLL]. For a semi-reduced basis the same proof can be applied.

**Lemma 5.2.2** The length of the shortest vector in a reduced basis differs from the length of a shortest non-zero vector in the lattice by at most a factor of \(2^{m-1} \). The length of a shortest vector in a semi-reduced basis differs from the length of a shortest non-zero vector in the lattice by at most a factor of \(2^m\).
A semi-reduced basis can be computed faster than a reduced basis, so we will use this slightly weaker notion in the sequel.

The next theorem establishes the relationship between shortest vectors in a semi-reduced basis and representations of algebraic numbers.

**Theorem 5.2.3** Let \( \mathbb{Q}(\alpha) \) be an algebraic number field, where \( \alpha \) is an algebraic integer with minimal polynomial \( p(X) = \sum_{i=0}^{n} p_i X^i, p_n = 1, p_i \in \mathbb{Z} \). Let \( \gamma \) be an element in \( \mathbb{Q}(\alpha) \) such that \( [\gamma] < K, K \geq 2 \). Assume \( s \) and \( \epsilon \) are real numbers satisfying

\[
s > 2^{2n^22^7n^n|p|_2^nK^{4n}, \ \epsilon = 4s^{-1}.\]

Moreover, suppose \( \bar{\gamma}, \bar{\alpha} \) are approximations to \( \gamma \) and \( \alpha \), respectively, such that the following estimates hold for the real and imaginary parts \( \Re, \Im \) of \( \gamma, \alpha \):

\[
|\Re(\gamma) - \Re(\bar{\gamma})| < \frac{1}{2} \epsilon, \ |\Im(\gamma) - \Im(\bar{\gamma})| < \frac{1}{2} \epsilon, \\
|\Re(\alpha_i) - \Re(\bar{\alpha}_i)| < \frac{1}{2} \epsilon, \ |\Im(\alpha_i) - \Im(\bar{\alpha}_i)| < \frac{1}{2} \epsilon, \ \forall i \in \{1, 2, \ldots, n-1\}.
\]

If \( \Lambda(V) \) is generated by the columns of the following \((n+3) \times (n+1)\) matrix

\[
V = \begin{bmatrix}
    s\Re(\gamma) & s\Re(\bar{\gamma}) & s\Re(\bar{\alpha}^2) & \cdots & s\Re(\bar{\alpha}^{n-1}) \\
    s\Im(\gamma) & 0 & s\Im(\bar{\alpha}^2) & \cdots & s\Im(\bar{\alpha}^{n-1}) \\
    1 & 0 & \cdots & & 0 \\
    0 & 1 & 0 & \cdots & 0 \\
    \vdots & & & & \vdots \\
    0 & 0 & \cdots & & 1
\end{bmatrix},
\]

then the shortest vector \( g = V(c, c_0, c_1, \ldots, c_{n-1})^T \) of a semi-reduced basis of \( \Lambda(V) \) satisfies

\[
\gamma = \frac{-1}{c} \sum_{i=0}^{n-1} c_i \alpha_i.
\]

Moreover, \( \text{gcd}(c, c_0, c_1, \ldots, c_{n-1}) = 1 \).

**Proof:** The columns in \( V \) are linearly independent. Each vector \( v \in \Lambda(V) \) can be identified with a unique vector \( (z, z_0, z_1, \ldots, z_{n-1}) \in \mathbb{Z}^{n+1} \) such that \( v = V(z, z_0, z_1, \ldots, z_{n-1})^T \). Every vector \( (z, z_0, z_1, \ldots, z_{n-1}) \in \mathbb{Z}^{n+1} \) in turn can be identified with a unique polynomial \( v(X, Y) = zX + \sum_{i=0}^{n-1} z_i Y^i \) in the
two variables $X$ and $Y$. Hence there is a one-to-one correspondence between vectors $\mathbf{v} \in \Lambda(V)$ and certain polynomials $v(X, Y) \in \mathbb{Z}[X, Y]$.

It is easily verified that the Euclidean norm $\|\mathbf{v}\|_2$ of a vector $\mathbf{v} \in \Lambda(V)$ satisfies

$$\|\mathbf{v}\|_2^2 = s^2|v(\gamma, \alpha)|^2 + |z|^2 + \sum_{i=0}^{n-1} |z_i|^2.$$  

Consider the vector $g = V(c, -c_0, \ldots, -c_{n-1})^T$ and the polynomial $g(X, Y) = cX - \sum_{i=0}^{n-1} c_i Y^i$, corresponding to the representation $\gamma = \frac{1}{n} \sum_{i=0}^{n-1} c_i \alpha^i$, $|c|, |c_i| < K$, $c, c_i \in \mathbb{Z}$. $g(\gamma, \alpha) = 0$ hence

$$|g(\gamma, \alpha)| = |g(\gamma, \alpha) - g(\gamma, \alpha)| \leq$$

$$\leq |c||\gamma - \gamma| + \sum_{i=1}^{n-1} |c_i||\alpha^i - \alpha^i| \leq nK\epsilon,$$

since by assumption $|c| < K$, $|c_i| < K$ for all $i = 0, \ldots, n - 1$. So the length $\ell$ of a shortest vector in a semi-reduced basis for $\Lambda(V)$ satisfies (see Lemma 5.2.2)

$$\ell \leq 2^{n+1}\|g\|_2 = 2^{n+1} \left( s^2|g(\gamma, \alpha)|^2 + |c|^2 + \sum_{i=0}^{n-1} |c_i|^2 \right)^{\frac{1}{2}} \leq$$

$$\leq 2^{n+1}((\epsilon sK)^2 + (nK)^2)^{\frac{1}{2}} \leq 2^{n+1}(\epsilon s + 1)nK.$$

By choice of $\epsilon$ and $s$

$$\ell < 2^{2n+3}K.$$

Next we claim that for any vector $\mathbf{v} \in \Lambda(V)$ whose Euclidean norm $\|\mathbf{v}\|_2$ is smaller than $2^{2n+3}K$ the corresponding polynomial $v(X, Y)$ satisfies $v(\gamma, \alpha) = 0$.

To prove the claim first note that

$$\|\mathbf{v}\|_2 = \left( s^2|v(\gamma, \alpha)|^2 + |z|^2 + \sum_{i=0}^{n-1} |z_i|^2 \right)^{\frac{1}{2}} \leq 2^{2n+3}K$$

implies

$$|v(\gamma, \alpha)| \leq 2^{2n+3}Ks^{-1}$$ and
\[ |z| \leq 2^{2n+3}K, \quad |z_i| \leq 2^{2n+3}K, \quad i = 0, \ldots, n - 1. \]

From \( |v(\gamma, \alpha)| \leq 2^{2n+3}Ks^{-1} \) we deduce
\[ |v(\gamma, \alpha)| \leq |v(\gamma, \alpha) - v(\overline{\gamma}, \overline{\alpha})| + |v(\overline{\gamma}, \overline{\alpha})| \leq 2^{2n+3}nK + 2^{2n+3}Ks^{-1} = 2^{2n+3}Ks^{-1}(4n + 1) < 2^{6n}Ks^{-1}, \]
where the bound on \( |v(\gamma, \alpha) - v(\overline{\gamma}, \overline{\alpha})| \) follows in exactly the same way as the corresponding bound for \( g \) shown above.

By assumption on the representation size of \( \gamma \) a non-zero integer \( c, |c| < K \), exists such that \( c\gamma \in \mathbb{Z}[\alpha] \subset \mathbb{R}[\alpha] \).

Consider the norm \( \text{no}(cv(\gamma, \alpha)) = \prod_{j=0}^{n-1} cv(\sigma_j(\gamma), \sigma_j(\alpha)) \) of \( cv(\gamma, \alpha) \). Since \( cv(\gamma, \alpha) \in \mathbb{R}[\alpha] \) this is a rational integer\(^9\). Hence
\[ \left| c^n \prod_{j=0}^{n-1} v(\sigma_j(\gamma), \sigma_j(\alpha)) \right| \in \mathbb{N} \cup \{0\}. \]

We show that the product is smaller than 1 and hence must be zero.
\[ \left| c^n \prod_{j=0}^{n-1} v(\sigma_j(\gamma), \sigma_j(\alpha)) \right| = |c^n||v(\gamma, \alpha)| \prod_{j=1}^{n-1} |v(\sigma_j(\gamma), \sigma_j(\alpha))| < s^{-1}2^{6n}K^{n+1} \prod_{j=1}^{n-1} |v(\sigma_j(\gamma), \sigma_j(\alpha))| \leq s^{-1}2^{6n}K^{n+1} \prod_{j=1}^{n-1} \left( |z\sigma_j(\gamma)| + \sum_{i=0}^{n-1} |z_i\sigma_j(\alpha)^i| \right). \]

Using the above estimates for \( |z| \) and \( |z_i| \) this shows
\[ \left| c^n \prod_{j=0}^{n-1} v(\sigma_j(\gamma), \sigma_j(\alpha)) \right| < s^{-1}2^n2^7nK^{2n} \prod_{j=1}^{n-1} \left( |\sigma_j(\gamma)| + \sum_{i=0}^{n-1} |\sigma_j(\alpha)^i| \right). \]

Writing \( \sigma_j(\gamma) \) as \( \frac{1}{7} \sum_{i=0}^{n-1} c_i \sigma_j(\alpha)^i \), combining corresponding powers of \( \sigma_j(\alpha) \), and expanding the product yields \( n^{(n-1)} \) terms each of which is smaller than \( (K + 1)^{n-1}M(p)^{n-1} \), where \( M(p) \) is the measure of \( p \) (see Definition 5.1.1, p. 48).

---

\(^9\)Recall from Equation (2) in Section 2 that the norm of an element in \( \mathbb{Q}(\alpha) \) corresponds to an integer power of the constant term of the minimal polynomial.
Hence by applying Landau’s bound for the measure of a polynomial (see Lemma 5.1.2, p. 48)

\[
\left| c^n \prod_{j=0}^{n-1} v(\sigma_j(\gamma), \sigma_j(\alpha)) \right| < s^{-1}2^{2n^2+7n}n^{n-1}K^{4n}|p|^{n-1}.
\]

By choice of \(s\) the claim follows.

Since \(c \neq 0\) \(\left| c^n \prod_{j=0}^{n-1} v(\sigma_j(\gamma), \sigma_j(\alpha)) \right| = 0\) implies that one of the factors in \(\prod_{j=0}^{n-1} v(\sigma_j(\gamma), \sigma_j(\alpha))\) must be zero. But these factors are conjugates of each other. Hence if one of them is zero all are, which proves our claim that \(\|v\| < 2^{2n+3}K\) implies \(v(\gamma, \alpha) = 0\).

Applying this result to the shortest vector \(g\) in a semi-reduced basis of the lattice \(\Lambda(V)\) shows that \(g\) corresponds to a polynomial \(g(X,Y) = cX + \sum_{i=0}^{n-1} c_i Y^i\) such that \(g(\gamma, \alpha) = 0\). This proves the first part of the theorem.

To prove the second claim observe that if \(\gcd(c, c_0, \ldots, c_{n-1}) \neq 1\) we would find a vector \(g' = V(c', c'_0, \ldots, c'_{n-1}) \in \Lambda(V)\) such that \(\gcd(c', c'_0, \ldots, c'_{n-1}) = 1\) and \(g'(\gamma, \alpha) = 0\). The unique representation of elements of \(Q(\alpha)\) as rational linear combinations of powers of \(\alpha\) shows \(g' = \frac{c}{c'} g\). Since \(\gcd(c', c'_0, \ldots, c'_{n-1}) = 1\) the integer \(c\) cannot properly divide \(c'\). On the other hand, \(g'\) can be written as an integer linear combination of the vectors in the semi-reduced basis. Since the elements of the basis are linearly independent even over \(Q\), this representation must be \(g' = \frac{c}{c'} g\). But then \(\frac{c'}{c} = +1\) or \(\frac{c'}{c} = -1\) which proves the second claim of the theorem.

The proof above is based on ideas of Lovász [Lo] but replaces a brute force estimate by Landau’s bound on the measure of a polynomial. Applying this modification to Lovász’ analysis of the corresponding bounds for minimal polynomials leads to an improvement of his bounds by a factor of \(n\). Thus this simple analysis can be used to deduce the same bounds as in [KLL].

Remark 5.2.4 The proof of the previous theorem can be generalized to give a quite simple analysis of an algorithm to reconstruct minimal polynomials over algebraic number fields and to factor polynomials over algebraic number fields (see also [La1],[Le]).

To turn Theorem 5.2.3 into an algorithm we need to compute a semi-
reduced basis for the lattice \( \Lambda(V) \). For lattices corresponding to matrices as in the theorem above the best run times are due to Schönhage [Sc4] although in general Schnorr’s algorithm [Scr] is better.

**Theorem 5.2.5 (Schönhage)** Let \( \alpha, \gamma, p, K, s, \epsilon \) be as in Theorem 5.2.3. A semi-reduced basis for the lattice generated by the columns of

\[
V = \begin{bmatrix}
  s \Re(\gamma) & s \Re(\pi) & s \Re(\pi^2) & \ldots & s \Re(\pi^{n-1}) \\
  s \Im(\gamma) & 0 & s \Im(\pi) & \ldots & s \Im(\pi^{n-1}) \\
  1 & 0 & \ldots & 0 \\
  0 & 1 & 0 & \ldots & 0 \\
  \vdots & \ddots & \ddots & \ddots & \ddots \\
  0 & 0 & \ldots & 0 & 1
\end{bmatrix},
\]

can be computed using \( O(n^2 \log s) \) elementary operations on integers whose binary length is bounded by \( O(\log s) \).

Observe that by choice of \( s \) and \( \epsilon \) we can assume that the entries of the matrix are integers.

Combining this result with Theorem 5.2.3 and interpreting everything in terms of bits yields

**Theorem 5.2.6** Let \( \mathbb{Q}(\alpha) \) be an algebraic number field, where \( \alpha \) is an algebraic integer with minimal polynomial \( p(X) = \sum_{i=0}^{n} p_i X^i, p_n = 1, p_i \in \mathbb{Z}, |p|_2 < 2^l \). Let \( \gamma \) be an element in \( \mathbb{Q}(\alpha) \) such that \( |\gamma| < 2^B \). Assume \( \epsilon > 0 \) satisfies

\[
\frac{1}{\epsilon} > 2n^2 + 7n + n \log n + nl + 4nB.
\]

Moreover, suppose that approximations \( \overline{\gamma}, \overline{\alpha} \) to \( \gamma \) and \( \alpha \), respectively, are given such that the following estimates hold for the real and imaginary parts \( \Re, \Im \) of \( \gamma, \alpha \)

\[
|\Re(\gamma) - \Re(\overline{\gamma})| < \epsilon, \quad |\Im(\gamma) - \Im(\overline{\gamma})| < \epsilon,
\]

\[
|\Re(\alpha^i) - \Re(\overline{\alpha}^i)| < \epsilon, \quad |\Im(\alpha^i) - \Im(\overline{\alpha}^i)| < \epsilon, \quad \forall i \in \{1, 2, \ldots, n-1\}.
\]

Then the representation of \( \gamma \) as \( \gamma = \frac{1}{\epsilon} \sum_{i=0}^{n-1} c_i \alpha^i, c, c_i \in \mathbb{Z}, \) and \( \gcd(c, c_0, c_1, \ldots, c_{n-1}) = 1 \), can be computed with \( O(n^3(n+l+B)) \) elementary operations on integers whose bit size is bounded by \( O(n(n+l+B)) \).

---

\[10\] Observe that by choice of \( s \) and \( \epsilon \) we can assume that the entries of \( V \) are integers.
Observe that the approximation to $\alpha$ we use in this theorem is much better than the one we require for the isolating rectangle (see the separation bound of Lemma 5.1.3, p. 49). Also observe that by the second claim in Theorem 5.2.3 the representation for $\gamma$ computed by the lattice basis reduction is normalized in the sense that the gcd of the integers appearing in the representation is 1. Therefore the integers of the representation will be smaller than $2^B$ in absolute value.

As we mentioned the technique we just presented to reconstruct algebraic numbers was originally invented by Kannan, Lenstra, Lovász [KLL] and (slightly different) Schönhage [Sc4] to reconstruct minimal polynomials from approximations. For later use let us state the precise result of Schönhage who has slightly better run times.

**Theorem 5.2.7 (Schönhage)** Let $\alpha$ be an algebraic integer such that the minimal polynomial $p$ of $\alpha$ is of degree at most $n$ and satisfies $|p|_2 \leq 2^l$. Given an approximation $\overline{\alpha}$ to $\alpha$ with $|\alpha - \overline{\alpha}| \leq 2^{-3(n^2 + n \log n + n + n l)}$ then the minimal polynomial $p$ can be reconstructed exactly using at most $O(n^3(n^2 + n l))$ elementary operations on integers of length at most $O(n^2 + n l)$.
5.3 Ratios of Radicals in Algebraic Number Fields

In this subsection we apply the result of the previous section to ratios of radicals in real and certain complex algebraic number fields. That is, we describe an algorithm which, given a "good" approximation to a ratio \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \) of radicals \( \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2} \) over an algebraic number field \( \mathbb{Q}(\alpha) \), computes an element \( \gamma \in \mathbb{Q}(\alpha) \) such that \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} = \gamma \) if the ratio is in \( \mathbb{Q}(\alpha) \). We derive a similar result for simple radicals \( \sqrt[d]{\rho} \).

**Lemma 5.3.1** Assume \( \sqrt[d]{\rho} \) is a radical over the number field \( \mathbb{Q}(\alpha) \), where \( \alpha \) is an algebraic integer with minimal polynomial \( p(X) \). If \( \sqrt[d]{\rho} \in \mathbb{Q}(\alpha) \) then

\[
[d^{1/2} \rho] < 3[\rho]^{2^n}|p|^{3^n}
\]

**Proof:** \( \sqrt[d]{\rho} \) is a solution to the equation \( X^d - \rho = 0 \). Let \( b \) be the denominator of \( \rho \). \( b\sqrt[d]{\rho} \) is a solution to \( X^d - b^d \rho = 0 \). This shows that \( b\sqrt[d]{\rho} \) is an algebraic integer (\( b^d \rho \) is an algebraic integer). Since we want to apply Lemma 5.1.8 we need a bound for \( [\sqrt[d]{\rho}]_{\infty} \).

The conjugates of \( \sqrt[d]{\rho} \) over \( \mathbb{Q} \) are among the \( d \)-th roots of the conjugates of \( \rho \). In fact, if \( f \) is the minimal polynomial of \( \rho \) then the polynomial \( f(X^d) \) has root \( \sqrt[d]{\rho} \). Therefore the conjugates of \( \sqrt[d]{\rho} \) must be among the roots of \( f(X^d) \), which in turn are all \( d \)-th roots of the conjugates of \( \rho \). Hence

\[
[\sqrt[d]{\rho}]_{\infty} \leq [\rho]^{1/2} \leq \max\{1, [\rho]_{\infty}\}
\]

and

\[
[b\sqrt[d]{\rho}]_{\infty} \leq b \max\{1, [\rho]_{\infty}\}.
\]

Here \( [\rho]^{1/2} \) denotes of course the positive real root.

By Lemma 5.1.11, p. 53, \( [\rho]_{\infty} < 3[\rho]|p|^{n/2} \). Therefore by Lemma 5.1.8, p. 51,

\[
[b\sqrt[d]{\rho}] < 3n^{2^n}|p|^{3^n}|\rho|^2.
\]

By Lemma 2.10, p. 22, the denominator of \( b\sqrt[d]{\rho} \in \mathbb{Q}(\alpha) \) is even bounded by \( |\Delta| \), where \( \Delta \) is the discriminant of \( \alpha \) (see Lemma 2.10, p. 22). Therefore the denominator of \( \sqrt[d]{\rho} \) is bounded by \( |b||\Delta| \). Hence the result follows from the bound on \( |\Delta|^{1/2} \) (see Lemma 5.1.5, p. 50).

If we interpret the lemma in terms of bit complexity it states that if \( |p|_2 < 2^l \) and \( |\rho| < 2^L \) then \( \sqrt[d]{\rho} \in \mathbb{Q}(\alpha) \) implies

\[
\log |\sqrt[d]{\rho}| < 2n \log n + 3nl + 2L + \log 3.
\]
For the rest of Section 5 define \( \tilde{L} := \lceil n \log n + nl + L \rceil \). Then the coefficient size above is strictly less than \( 2^{3\tilde{L}} \).

The important thing to notice here is that the bounds are independent of \( d \). This may seem quite surprising but it only reflects the fact that the size of the coefficients of factors of a polynomial depend only on the degree of the factors and on the size of the coefficients of the polynomial but not on its degree (see [WR]).

Next we deduce a similar bound for ratios of radicals over real algebraic number fields or certain complex algebraic number fields.

**Lemma 5.3.2** Suppose that \( \mathbb{Q}(\alpha) \) is either a real algebraic number field or a number field containing \( d_1 \)-th and \( d_2 \)-th primitive roots of unity. In both cases let \( \alpha, p \) be as above. If \( \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2} \) are either real radicals in the real case or arbitrary radicals in the complex case then \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \in \mathbb{Q}(\alpha) \) implies

\[
\left[ \frac{\sqrt[d_1]{\rho_1}}{\sqrt[d_2]{\rho_2}} \right] < 3n^{6n} |\rho_1|^{5n} |\rho_1|^{2} |\rho_2|^{2n}.
\]

**Proof:** First we prove a bound on the denominator of \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \).

By Lemma 3.4, p.24, \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \in \mathbb{Q}(\alpha) \) implies \( \sqrt[d_1]{\rho_1}d \in \mathbb{Q}(\alpha) \), \( \sqrt[d_2]{\rho_2}d \in \mathbb{Q}(\alpha) \), where \( d = \gcd(d_1, d_2) \). Furthermore let \( d_1 = d_1/d, d_2 = d_2/d \).

Hence \( \sqrt[d_1]{\rho_1}d \) is a root of \( X^{d_1} - \rho_1 \) and \( \sqrt[d_2]{\rho_2}d \) is a root of \( X^{d_2} - 1/\rho_2 \).

By the same argument as in the proof of Lemma 5.3.1 \( \sqrt[d_1]{\rho_1}d \) is an algebraic integer for some \( b \in \mathbb{Z} \), \( b \leq [\rho_1] \).

We need a similar result for \( \sqrt[d_2]{\rho_2}d \). Again we only need to bound the size of a rational integer \( b' \) such that \( b' \sqrt[d_2]{\rho_2}d \) is an algebraic integer.

As has been observed in Section 2 if \( \sqrt[d_2]{\rho_2}d \) is a root of the polynomial \( r \in \mathbb{Z}[X] \) whose leading coefficient is \( r_m \) then \( r_m \sqrt[d_2]{\rho_2}d \) is an algebraic integer.

By Lemma 5.1.11 \( \sqrt[d_2]{\rho_2}d \) is root of a polynomial \( r \) whose 2-norm \( |r|_2 \) is bounded by \( n^{2n} |\rho_2|^{[\rho_2]^{2}} \). This gives us the bound on \( b' \). Moreover, combined with
bound on \( b \) it shows that an integer \( c \) exists such that

\[
c \left( \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \right)^d \in R_\alpha, \quad |c| < |p|^{2n} [p_1][p_2]^n.
\]

Next observe that \( \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \) is a root of

\[
X^d - \left( \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \right)^d.
\]

Using again the argument of the proof of the previous lemma it has been shown that if \( \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \in \mathbb{Q}(\alpha) \) then an integer \( c < |p|^{2n}|p_1|[p_2]^n \) exists such that

\[
c \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \in R_\alpha.
\]

By Lemma 5.1.8, p.51, and Lemma 5.1.7, p. 51,

\[
\left[ \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \right] < 3n^2|p|^{2n}|c||p_1|[p_2]^{-1}|\alpha\|_\infty.
\]

By Lemma 2.10, p. 22, the denominator of \( \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \) is bounded by \( \Delta \), where \( \Delta \) is the discriminant of \( \alpha \). Hence the denominator of \( \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \) is bounded by \( |c||\Delta| \). Using the bound on \( c \) from above, the bound for \( \Delta \) (Lemma 5.1.5, p.50), and the bounds for \( |p_1|, |p_2|^{-1} \) from Lemma 5.1.11, p.53, the lemma follows.

If we assume \( |p|_2 < 2^l \) and \( |p_1|, |p_2| < 2^L \) then

\[
\log \left[ \frac{\sqrt[4]{p_1}}{\sqrt[4]{p_2}} \right] < 6n \log n + 5nl + 3nL + \log 3.
\]

Defining for the rest of Section 5 \( \mathcal{L} := \lfloor n \log n + nl + nL \rfloor \) the coefficient size of a ratio of radicals is strictly less than \( 2^{6 \mathcal{L}} \).

Using the bounds of the previous two lemmata and approximations to \( \alpha \) and \( \sqrt[p_1]{\alpha} \) or \( \sqrt[p_1]{\alpha} / \sqrt[p_2]{\alpha} \) we can apply Theorem 5.2.6 in order to determine the integer vector \((c, c_0, c_1, \ldots, c_{n-1}) \in \mathbb{Z}^{n+1}\) such that if \( \sqrt[p_1]{\alpha} \) or \( \sqrt[p_1]{\alpha} / \sqrt[p_2]{\alpha} \) is in \( \mathbb{Q}(\alpha) \) then its representation is \( \frac{1}{\ell} \sum_{i=0}^{n-1} c_i \alpha^i \).
Corollary 5.3.3 Suppose $\sqrt[d]{\rho}$ is a radical over an algebraic number field $\mathbb{Q}(\alpha)$. Here $\alpha$ is an algebraic integer with minimal polynomial $p(X) = \sum_{i=0}^{n} p_i X^i, p_n = 1, p_i \in \mathbb{Z}, |p|_2 < 2^l$. Assume $[\rho] < 2^L$ and let $\epsilon > 0$ be such that

$$\log \frac{1}{\epsilon} > 2n^2 + n \log n + 7n + nl + 12nL.$$ 

Moreover, suppose that approximations $\bar{\gamma}, \bar{\alpha}$ to $\gamma = \sqrt[d]{\rho}$ and $\alpha$, respectively, are given such that the following estimates for the real and imaginary parts $\Re, \Im$ of $\gamma, \alpha$ hold

$$|\Re(\gamma) - \Re(\bar{\gamma})| < \epsilon, \quad |\Im(\gamma) - \Im(\bar{\gamma})| < \epsilon,$$

$$|\Re(\alpha^i) - \Re(\bar{\alpha}^i)| < \epsilon, \quad |\Im(\alpha^i) - \Im(\bar{\alpha}^i)| < \epsilon, \quad \forall i \in \{1, 2, \ldots, n-1\}.$$ 

Then integers $c, c_i, \gcd(c, c_0, c_1, \ldots, c_{n-1}) = 1$, such that if $\sqrt[d]{\rho} \in \mathbb{Q}(\alpha)$ then

$$\sqrt[d]{\rho} = \frac{1}{e} \sum_{i=0}^{n-1} c_i \alpha^i$$

can be computed using $O(n^3 \tilde{L})$ elementary operations on integers of bit size at most $O(nL)$.

Observe that in this corollary we may even assume that $\sqrt[d]{\rho}$ is specified by the approximation rather than by the form given in the introduction. It should be noted however that the approximation does not necessarily specify a unique root of $\rho$.

This corollary can be used to check whether $\mathbb{Q}(\alpha)$ contains certain roots of unity. This test in turn can be used to check the condition for $\mathbb{Q}(\alpha)$ we need for the linear dependence result on complex radicals over $\mathbb{Q}(\alpha)$.

Replacing the bound from Lemma 5.3.1 by the bound in Lemma 5.3.2 we similarly get

Corollary 5.3.4 Suppose $\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}$ are either real radicals over a real algebraic number field $\mathbb{Q}(\alpha)$ or arbitrary radicals over a number field $\mathbb{Q}(\alpha)$ containing primitive $d_1$-th and $d_2$-th roots of unity. Here $\alpha$ is an algebraic integer with minimal polynomial $p(X) = \sum_{i=0}^{n} p_i X^i, p_n = 1, p_i \in \mathbb{Z}, |p|_2 < 2^l$. Let $[\rho_1], [\rho_2] < 2^L$ and assume that $\epsilon > 0$ satisfies

$$\log \frac{1}{\epsilon} > 2n^2 + n \log n + 7n + nl + 24nL.$$ 

Moreover, suppose that approximations $\bar{\gamma}, \bar{\alpha}$ to $\gamma = \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2}$ and $\alpha$, respectively, are given such that the following estimates hold

$$|\Re(\gamma) - \Re(\bar{\gamma})| < \epsilon, \quad |\Im(\gamma) - \Im(\bar{\gamma})| < \epsilon,$$
\[ |\Re(\pi^i) - \Re(\alpha^i)| < \epsilon, \ |\Im(\pi^i) - \Im(\alpha^i)| < \epsilon, \ \forall i \in \{1, 2, \ldots, n-1\}. \]

Then integers \( c, c_0, c_1, \ldots, c_{n-1}, \gcd(c, c_0, c_1, \ldots, c_{n-1}) = 1, \) such that if \( \sqrt[n]{\rho_1}/\sqrt[n]{\rho_2} \in \mathbb{Q}(\alpha) \) then \( \sqrt[n]{\rho_1}/\sqrt[n]{\rho_2} = \frac{1}{\epsilon} \sum_{i=0}^{n-1} c_i \alpha^i \) can be computed exactly using \( \mathcal{O}(n^3 \mathcal{L}) \) elementary operations on integers of bit size at most \( \mathcal{O}(n\mathcal{L}) \).

Of course in the real case all imaginary parts are zero. As above, we may assume that the radicals are specified simply by the approximations.

The results in this section are restricted in two directions. First, the two corollaries above assume that approximations to radicals or ratios of radicals are already given. And second, the corollaries do not tell us whether a radical or a ratio of radicals is contained in some number, rather they tell us how to compute candidate representations for these numbers in the algebraic number field. The next two subsections show how to resolve these restrictions.
5.4 Approximating Radicals and Ratios of Radicals

In Corollary 5.3.3 and Corollary 5.3.4 we assumed that approximations to $\alpha$, $\sqrt[\rho]{\cdot}$, $\sqrt[1/\rho]{\cdot}$, respectively, are given. In this subsection we show how to compute efficiently such approximations. In particular, the run time of the approximation algorithms will depend polynomially only on $\log d_i$ rather than on $d_i$. Our algorithms use several well-known approximation algorithms. The first one is due to Schönhage [Sc3].

Theorem 5.4.1 (Schönhage) Let $p(X) \in \mathbb{Z}[X]$ be an integer polynomial such that $\deg p = n$, $|p| < 2^n$. Furthermore assume $\epsilon < 2^{-n\log n - n}$. Then approximations $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ to the roots $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ of $p$ with $|\alpha_i - \alpha_i| < \epsilon$ can be computed using $O(n)$ elementary operations on floating-point numbers of size at most $O(n \log \frac{1}{\epsilon})$.

We assume $\epsilon < 2^{-n\log n - n}$ because this bound suffices to separate the roots $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ of $p$ (see Lemma 4.2.1, p. 41). Hence the approximations $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ will be $n$ distinct complex numbers. Moreover, the output of this algorithms will be $n$ floating-point numbers of the kind we defined in the introduction to this thesis (see the introduction of [Sc3]). The second type of approximation algorithms is due to Brent [Br].

Theorem 5.4.2 (Brent) 1. $\pi$ can be computed with relative error less than $O(2^{-n})$ by $O(\log n)$ elementary operations on floating-point numbers of size at most $O(n)$.

2. Let $-\infty < a < b < \infty$. If $x$ is an $n$-bit floating-point number in $[a, b]$ then $\exp(x)$ can be computed with relative error less than $O(2^{-n})$ by $O(\log n)$ elementary operations on floating-point numbers of size $O(n)$. If $a > 0$ the same is true for $\ln x$.

3. For all $n$-bit floating-point numbers $x$ $\arctan(x)$ can be computed with relative error less than $O(2^{-n})$ by $O(\log n)$ elementary operations on floating-point numbers of size $O(n)$.

4. Let $-\pi < a < b < \pi$. If $x$ is an $n$-bit floating-point number in $[a, b]$ then $\sin(x)$ and $\cos(x)$ can be computed with relative error less than $O(2^{-n})$ using $O(\log n)$ elementary operations on floating-point numbers of size $O(n)$.

Since $\sin(x), \cos(x)$ are bounded in absolute value by 1 and $\arctan(x)$ is bounded in absolute error by $\pi$ Brent’s algorithms can actually be used to
compute the trigonometric functions with absolute error less than $O(2^{-n})$ within the time bounds stated above. Of course the same is true for $\pi$ and, more important, since $[a, b]$ is a fixed interval it is also correct for $\ln(x)$ and $\exp(x)$ for all $x$ in the interval $[a, b]$.

We also need several standard estimates on absolute errors. We summarize them in the following lemma.

**Lemma 5.4.3** Let $x, \overline{x}$ be a complex numbers satisfying $2^{-m} < |x| < 2^m$, $0 < \epsilon < 2^{-(m+1)}$, $|\overline{x} - x| < \epsilon$. Then

1. $|x^i - \overline{x}^i| < \epsilon 2^{(m+1)i}$.

2. If $\max\{|b_0|, |b_1|, \ldots, |b_{n-1}|\} < 2^L$ then

$$\left|\sum_{i=0}^{n-1} b_i x^i - \sum_{i=0}^{n-1} b_i \overline{x}^i\right| < \epsilon 2^{(m+1)n+L}.$$

3. $\left|\frac{1}{x} - \frac{1}{\overline{x}}\right| < \epsilon 2^{m+3}$.

4. If $d \geq 2$ then

$$\left|x^\frac{1}{d} - \overline{x}^\frac{1}{d}\right| < 2^{m+1}\epsilon.$$

The first two bounds follow from standard estimates. The third and fourth bound are simple consequences of the Mean Value Theorem in its complex version (see for example [C]).

The following lemma is the crucial step in the analysis of the approximation algorithms.

**Lemma 5.4.4** Assume $x \in \mathbb{R}$, $\epsilon$ satisfy the bounds of the previous lemma, $d, k \in \mathbb{N}$, and $d > 2$, $k < d$. If

$$y \in I = \left(1 - \epsilon \exp \left\{ (1 - \epsilon) \frac{k}{d} \ln x \right\}, (1 + \epsilon) \exp \left\{ (1 + \epsilon) \frac{k}{d} \ln x \right\}\right)$$

then $y$ is an approximation to $\exp(\frac{k}{d} \ln x)$, i.e. $x^\frac{k}{d}$ with relative error less than $4m\epsilon$. 

69
Proof: First we prove a lower bound on the left endpoint of the interval $I$.

\[
(1 - \epsilon) \exp \left\{ (1 - \epsilon) \frac{k}{d} \ln x \right\} = \\
= (1 - \epsilon) \exp \left( \frac{k}{d} \ln x \right) \exp \left( -\epsilon \left( \frac{k}{d} \ln x \right) \right) > \\
> x^\frac{k}{d} (1 - \epsilon) \left( 1 - \epsilon \frac{k}{d} \ln x \right),
\]

since $\exp(\delta) > 1 + \delta$ for all $\delta$.

Therefore

\[
(1 - \epsilon) \exp \left\{ (1 - \epsilon) \frac{k}{d} \ln x \right\} \ln x > (1 - 2m\epsilon)x^\frac{k}{d}.
\]

Similarly we get an upper bound for the right endpoint of $I$.

\[
(1 + \epsilon) \exp \left\{ (1 + \epsilon) \frac{k}{d} \ln x \right\} = \\
= (1 + \epsilon) \exp \left( \frac{k}{d} \ln x \right) \exp \left( \epsilon \left( \frac{k}{d} \ln x \right) \right) < \\
< x^\frac{k}{d} (1 + \epsilon) \left( 1 + 2\epsilon \frac{k}{d} \ln x \right),
\]

since $\exp \delta < 1 + 2\delta$ for $|\delta| \leq \frac{1}{2}$.

Hence

\[
(1 + \epsilon) \exp \left\{ (1 + \epsilon) \frac{k}{d} \ln x \right\} < (1 + 4m\epsilon)x^\frac{k}{d}.
\]

\[
\square
\]

Now we can analyze the time needed to compute the approximations required by Corollary 5.3.3 and Corollary 5.3.4. First let us state the result for approximating powers of $\alpha$.

Lemma 5.4.5 Let $\alpha$ be root of the polynomial $p(X) = \sum_{i=0}^{n} p_i x^i, p_n = 1, p_i \in \mathbb{Z}, |p_i| < 2^i$. An approximation $\overline{\alpha}$ to $\alpha$ satisfying $|\alpha^i - \overline{\alpha}^i| < \epsilon \leq 2^{-n \log n - n \epsilon}$ for all $i < n$, can be computed using $\mathcal{O}(n)$ elementary operations on floating-point numbers of bit size $\mathcal{O}(n \log \frac{1}{\epsilon})$. 

70
Proof: Set $\epsilon' := \epsilon 2^{-2(l+1)n}$. By Lemma 5.4.3 an approximation $\overline{\alpha}$ with $|\alpha - \overline{\alpha}| < \epsilon'$ will lead to approximations as required. By Schönhage's algorithm (Theorem 5.4.1) the lemma follows.

Lemma 5.4.6 Suppose $\sqrt[p_1]{\alpha}, \sqrt[p_2]{\alpha}$ are real radicals over a real algebraic number field $\mathbb{Q}(\alpha)$ where $\alpha$ is an algebraic integer with minimal polynomial $p(X) = \sum_{i=0}^{n} p_i X^i$, $p_n = 1, p_i \in \mathbb{Z}$, $|p|_2 < 2^l$, and let $L = \lceil n \log n + nl + nL \rceil$.

For any $\epsilon < 2^{-2L}$ approximations to $\sqrt[p_1]{\alpha}, \sqrt[p_2]{\alpha}$ with absolute error less than $\epsilon$ can be computed using $O(n)$ elementary operations on floating-point numbers of size $O(n \log \frac{1}{\epsilon})$, $O(\log \log \frac{1}{\epsilon})$ elementary operations on floating-point numbers of size $O(\log \frac{1}{\epsilon})$, and a constant number of operations on floating-point numbers of length $O(\log \frac{1}{\epsilon} + \max\{\log d_i\})$.

Within the same time bounds an approximation to the ratio $\sqrt[p_1]{\alpha} / \sqrt[p_2]{\alpha}$ with absolute error less than $\epsilon$ can be computed.

Proof: Of course we show the first statement only for $\sqrt[p_1]{\alpha}$. Furthermore since we are dealing with real radicals in this lemma we may assume that $\sqrt[p_1]{\alpha}$ is given by $\exp \left( \frac{1}{p_1} \right)$.

First we compute an approximation $\overline{p_1}$ to $p_1$ with absolute error less than $\epsilon_1 < \epsilon$, where $\epsilon_1$ will be specified later. This can be done as follows.

Assume $p_1 = \frac{1}{b} \sum_{i=0}^{n-1} b_i \alpha^i$, $b, b_i \in \mathbb{Z}$. Compute an approximation to $\frac{1}{b}$ with absolute error less than $\epsilon_1 2^{-(L+ln+4)}$. Since $\frac{1}{b} \leq 1$ an approximation with relative error $\epsilon_1 2^{-(L+ln+4)}$ suffices. This approximation can be determined by a constant number of elementary operations on floating-point numbers of size $O(\log \frac{1}{\epsilon_1})$.

Also compute an approximation to $\sum_{i=0}^{n-1} b_i \alpha^i$ with absolute error less than $\frac{1}{4}\epsilon_1$. If $\overline{\alpha}$ is an approximation to $\alpha$ satisfying $|\alpha - \overline{\alpha}| < \epsilon_1 2^{-(L+2(l+1)n+2)}$ then $\sum_{i=0}^{n-1} b_i \overline{\alpha}^i$ is such an approximation (see Lemma 5.4.3). The approximation $\overline{\alpha}$ can be computed with $O(n)$ elementary operations on floating-point numbers of size $O(n \log \frac{1}{\epsilon_1})$). It is easy to see that this dominates the time needed to compute $\sum_{i=0}^{n-1} b_i \overline{\alpha}^i$. Finally multiplying the approximations gives the result since $\frac{1}{b} \leq 1$ and $\sum_{i=0}^{n-1} b_i \alpha^i < 2^{L+ln+2}$ (see Lemma 5.1.11, p. 53).

Also by Lemma 5.1.11 and by the choice of $\epsilon$ this approximation suffices to determine the sign of $p_1$, hence for the rest of the proof we can assume that $p_1 > 0$. Moreover, the approximation $\overline{p_1}$ will be non-zero which is important for the proof of the second part of the lemma.
Next an approximation $\gamma_1$ to $\frac{1}{\sqrt[4]{\rho_1}}$ with absolute error less than $\epsilon_1$ is computed in the following way.

$\rho_1$ is given in the form $a2^m$ for some $a, m \in \mathbb{Z}$ such that $\log |a|, |m| \leq \frac{1}{\epsilon_1}$ (Lemma 5.1.11). Write $\rho_1$ as $a2^{m_1}2^{m_2}$ with $a2^{m_1} \in [2, 4)$. This can be done with $|m_1| \leq \log \frac{1}{\epsilon_1}$, $m_2 \leq L$. Then

$$\rho_1^{\frac{1}{4}} = (a2^{m_1})^{\frac{1}{4}} 2^{\frac{m_2}{4}},$$

taking positive roots if $d_1$ is even.

We determine the representation of $m_2$ as $ud_1 + k$ with $u, k \in \mathbb{Z}$, $0 \leq k < d$. This can be done by a constant number of operations on numbers of size $O(\log \frac{1}{\epsilon_1} + \log d_1)$.

Now $\rho_1^{\frac{1}{4}} = (a2^{m_1})^{\frac{1}{4}} 2^{\frac{k}{4}}$. If $(a2^{m_1})^{\frac{1}{4}}$ and $2^{\frac{k}{4}}$ are both computed with absolute error less than $2^{-1(L+3)}\epsilon_1$, then the product of these approximations and $2^u$ leads to an approximation $\gamma_1$ as required. Obviously the product can be computed within the time bounds stated in the lemma. So it remains to analyze how much time is needed to determine the approximations to $(a2^{m_1})^{\frac{1}{4}}$ and $2^{\frac{k}{4}}$.

$$(a2^{m_1})^{\frac{1}{4}} = \exp \frac{1}{d_1} \ln a2^{m_1}$$

and

$$2^{\frac{k}{4}} = \exp \frac{k}{d_1} \ln 2.$$

By the previous lemma if $\frac{1}{d_1} \ln a2^{m_1}$ and $\frac{k}{d_1} \ln 2$ are computed with relative error less than $\epsilon_2$ and the exp of these values is approximated with relative error less than $\epsilon_2$ then this leads to approximations to $(a2^{m_1})^{\frac{1}{4}}$ and $2^{\frac{k}{4}}$ with relative error less than $8\epsilon_2$. But both numbers are bounded in absolute value by 2 so these approximations are with absolute error less than $16\epsilon_2$. Choosing $\epsilon_2 = 2^{-(L+7)}\epsilon_1$ therefore suffices to determine approximations to $(a2^{m_1})^{\frac{1}{4}}$, $2^{\frac{k}{4}}$ with absolute error less than $\epsilon_12^{-(L+3)}$, and, accordingly, an approximation to $\sqrt[4]{\rho_1}$ with absolute error less than $\epsilon_1$.

By Brent’s results (Theorem 5.4.2) we can compute the required approximations using $O(\log \log \frac{1}{\epsilon_1})$ operations on floating-point numbers of size $O(\log \frac{1}{\epsilon_1})$ and a constant number of operations on floating-point numbers of size $O(\log \frac{1}{\epsilon_1} + \log d_1)$. The last term is caused by computing the inverse of $d_1$ with precision $2^{-O\left(\log \frac{1}{\epsilon_1}\right)}$. 

72
Finally observe that by Lemma 5.1.11, p. 53, $|\sqrt[2]{\rho_1}| < \max\{1,|\rho_1|\} < 2^{n_l+L+2}$ and hence Lemma 5.4.3 shows

$$\frac{1}{\gamma_1} - \frac{1}{\gamma_1} \leq \left| \frac{1}{\gamma_1} - \frac{1}{\sqrt[2]{\rho_1}} \right| + \left| \frac{1}{\sqrt[2]{\rho_1}} - \frac{1}{\gamma_1} \right| < 2^{L+nl+3}\epsilon_1 + \epsilon_1 = 2^{L+nl+4}\epsilon_1.$$  

Choosing $\epsilon_1 = e^{-(L+nl+4)}$ therefore proves the first part of the lemma.

To prove the second claim of the lemma note that by Lemma 5.1.11, p. 53, not only $|\sqrt[2]{\rho_1}| < \max\{1,|\rho_1|\} < 2^{n_l+L+2}$ but also $|\sqrt[2]{\rho_2}| < 2^{L\epsilon_2}$. Computing $\frac{1}{\gamma_1}$ with absolute error $\epsilon_2 = e^{-(n_l+L+4)}$ and $\frac{1}{\sqrt[2]{\rho_2}}$ with absolute error $\epsilon_2 = e^{-(n_l+L+4)}$ therefore suffices to prove the claim. The time required to determine $\gamma_1$ is analyzed by the first part of the lemma.

To analyze the time needed for approximating $\frac{1}{\gamma_2}$ assume that $\gamma_2$ is an approximation to $\frac{1}{\sqrt[2]{\rho_2}}$ with $|\gamma_2 - \frac{1}{\sqrt[2]{\rho_2}}| < \epsilon_2 2^{-4(L+1)}$ and $\frac{1}{\gamma_2}$ is an approximation to $\frac{1}{\gamma_2}$ with absolute error less than $\frac{1}{2}\epsilon_2$ then

$$\left| \frac{1}{\gamma_2} - \frac{1}{\gamma_2} \right| \leq \left| \frac{1}{\gamma_2} - \frac{1}{\sqrt[2]{\rho_2}} \right| + \left| \frac{1}{\sqrt[2]{\rho_2}} - \frac{1}{\gamma_2} \right| \leq$$

$$\frac{1}{2}\epsilon_2 + \frac{1}{2}\epsilon_2 \leq \epsilon_2,$$

since the bound on $|\sqrt[2]{\rho_2}|$ and Lemma 5.4.3 show $\left| \frac{1}{\gamma_2} - \frac{1}{\sqrt[2]{\rho_2}} \right| \leq \frac{1}{2}\epsilon_2$.

By the first part of the lemma $\gamma_2$ can be determined within the time bounds stated.

Since $|\gamma_2 - \frac{1}{\sqrt[2]{\rho_2}}| < \epsilon_2 2^{-4(L+1)}$ and $|\sqrt[2]{\rho_2}| > 2^{-2L}$ the approximation $\gamma_2$ satisfies $|\gamma_2| > 2^{-2L+1}$. Therefore an approximation to $\frac{1}{\gamma_2}$ with relative error less than $\epsilon_2 2^{-2L+2}$ leads to the approximation $\frac{1}{\gamma_2}$. The lemma follows.

Observe that the approximations to $\alpha$ may also be computed by successive bisecting the isolating interval for $\alpha$ and using the theory of Sturm sequences (see [vdW]) to determine after a bisection step which of the two created intervals contains $\alpha$. Using Schönhages result is at least theoretically more efficient.

**Corollary 5.4.7** In the real case approximations as required by Corollary 5.3.4 can be computed using $O(n)$ elementary operations on floating-point

\[\text{Note that all denominators are non-zero.}\]
numbers of size $O(n^2L)$, $O(\log nL)$ elementary operations on floating-point
numbers of length $O(nL)$ plus a constant number of operations on floating-
point numbers of length $O(nL + \max\{\log d_1, \log d_2\})$.

**Proof:** The $\epsilon$ in Corollary 5.3.4 satisfies $\log \frac{1}{\epsilon} = O(nL)$.

Hence in time polynomial in $n, l, L, \log d_1, \log d_2$ we can compute approxi-
mations to ratios of real radicals as required by the lattice basis reduction
algorithm in Corollary 5.3.4. Observe that by the algorithm leading to
Lemma 5.4.6 we can efficiently compute the sign of $\rho_i$ so from now on the
assumption that certain radicals $\sqrt[\zeta_d]{\rho_i}$ are real is justified since it can be
checked efficiently.

Next we show the corresponding results for approximations to complex
radicals. Recall from the introduction to this section that we assume that
the radical $\sqrt[\zeta_d]{\rho}$ is given by

$$\zeta_d \sqrt[\zeta_d]{\rho} = \zeta_d |\rho|^{\frac{1}{d}} (\cos \frac{1}{d} \phi_d + i \sin \frac{1}{d} \phi_d),$$

where $\phi_d$ denotes the
angle of $\rho$, $\zeta_d = \cos \frac{2\pi d}{d} + i \sin \frac{2\pi d}{d}$, and $0 \leq k < d$.

To give a rigorous analysis of the run times of the approximation algo-
rithm for complex radicals of this form we need the following lemma.

**Lemma 5.4.8** Let $\rho$ be a root of $f = \sum_{i=0}^m f_i X^i \in \mathbb{Z}[X], \ |f|_2 < 2^h$. If $\Re(\rho), \Im(\rho) \neq 0$ then $|\Re(\rho)| > 2^{-(m \log m + mh + 1)}$ and $|\Im(\rho)| > 2^{-(m \log m + mh + 1)}$.

**Proof:** $|\Im(\rho)| = \frac{1}{2} |\rho - \bar{\rho}|$, where $\bar{\rho}$ is the complex conjugate of $\rho$. $\bar{\rho}$ is also a
root of $f$. Hence the root separation bound (see Lemma 5.1.3, p. 49) applies. To get the bound on $\Re$ apply the same argument to $i\rho$ and the polynomial $F(\frac{1}{2}X) \in \mathbb{Z}[i][X]$. Although $F(\frac{1}{2}X)$ is not an integer polynomial the root
separation bound still applies (see for example [Ja] and [Mi1]).

**Lemma 5.4.9** Suppose $\mathbb{Q}(\alpha)$ is a number field generated by the algebraic
integer $\alpha$, whose minimal polynomial is $p(X) \in \mathbb{Z}[X], \ |p|_2 < 2^l$. If $d_1, d_2 \in \mathbb{N}$
and $\rho_1, \rho_2 \in \mathbb{Q}(\alpha)$ satisfy $|\rho_i| < 2^l$ then for any $\epsilon < 2^{-2nL-n \log n-1}$
the radicals $\sqrt[\zeta_d]{\rho_i}, i = 1, 2$, can be computed with absolute error less
than $\epsilon$ using $O(n)$ elementary operations on floating-point numbers of size
$O(n \log \frac{1}{\epsilon})$, $O(\log \log \frac{1}{\epsilon})$ elementary operations on floating-point numbers
of size $O(\log \frac{1}{\epsilon})$, and a constant number of operations on floating-point numbers
of bit size $O(\max\{\log d_i\} + \log \frac{1}{\epsilon})$.

Within the same time bounds the ratio of these radicals can be computed
with absolute error less than $\epsilon$.
Proof: As in the proof of Lemma 5.4.6 we show the first statement only for \( \sqrt[2d_1]{\rho_1} \).

First within the time bounds stated \(|\rho_1|^{\frac{1}{2d_1}}\) is approximated with absolute error less than \( \frac{1}{2} \epsilon \). This is done in the following fashion. Suppose \( \overline{\rho_1} \) is an approximation to \( \rho_1 \) with absolute error less than \( \epsilon 2^{-(nl+L+6)} \). \(|\rho_1 - \overline{\rho_1}| < \epsilon 2^{-(nl+L+6)} \) implies

\[
||\rho_1| - |\overline{\rho_1}||| < \epsilon 2^{-(nl+L+6)}.
\]

Therefore (see Lemma 5.4.3 and Lemma 5.1.11, p. 53 for a bound on \(|\rho_1|\))

\[
\left| |\overline{\rho_1}|^{\frac{1}{2d_1}} - |\rho_1|^{\frac{1}{2d_1}} \right| < \frac{1}{8} \epsilon.
\]

Hence determining \( |\overline{\rho_1}|^{\frac{1}{2d_1}} = \left( \sqrt{\Re(\rho_1)^2 + \Im(\rho_1)^2} \right)^{\frac{1}{2d_1}} = 2d_1 \sqrt{\Re(\rho_1)^2 + \Im(\rho_1)^2} \) with absolute error less than \( \frac{1}{2} \epsilon \) leads to an approximation to \( |\rho_1|^{\frac{1}{2d_1}} \) as desired. Computing the \( 2d_1 \)-th root can be analyzed exactly as in Lemma 5.4.6 so we skip the details.

Next we show how to compute \( \cos \frac{1}{2d_1} \phi_1 + i \sin \frac{1}{2d_1} \phi_1 \) and \( \zeta^k = \cos \frac{2k\pi}{d_1} + i \sin \frac{2k\pi}{d_1} \) with relative and, since the absolute value of \( \cos \frac{1}{2d_1} \phi_1 + i \sin \frac{1}{2d_1} \phi_1 \) is equal to 1, with absolute error less than

\[
\epsilon_1 = \epsilon 2^{-(nl+L+6)}.
\]

Since \(|\rho_1| < 2^{nl+L+2}\) (see Lemma 5.1.11, p. 53) this suffices to show that the product of the approximated values of \( \cos \frac{1}{2} \phi_1 + i \sin \frac{1}{2} \phi_1 \), \( \zeta^k_1 \), and of \( |\rho_1|^{\frac{1}{2d_1}} \) will lead to an approximation to \( \sqrt[2d_1]{\rho_1} \) with absolute error less than \( \epsilon \).

First observe that combining the previous lemma with the bounds of Lemma 5.1.11, p. 53, shows \(|\Re(\rho_1)|, |\Im(\rho_1)| > 2^{-2nL-n\log n} - 1\) if these values are non-zero. Therefore an approximation to \( \rho_1 \) with absolute error less than \( \epsilon_1 < \epsilon \) are sufficient to determine whether the real or imaginary part are non-zero. If any of them is non-zero the approximation also allows us to determine the sign of the real and imaginary part of \( \rho_1 \). In particular, If the real part is non-zero the approximation to the real part will be non-zero, too. Hence we can determine the quadrant containing the angle \( \phi_{\rho_1} \). So for the rest of the proof assume \( 0 \leq \phi_{\rho_1} \leq \frac{\pi}{2} \).

We show how to approximate \( \phi_{\rho_1} \). If the real or the imaginary part is zero this is trivial so we assume that both are non-zero. Next observe \( \phi_{\rho_1} = \arctan \left( \frac{\Im(\rho_1)}{\Re(\rho_1)} \right) \). Because of the lower bound on \(|\Re(\rho_1)|\) of the previous lemma within the time bounds stated in the lemma an approximation \( \rho \)
to \( \frac{\Re(\rho_1)}{\Im(\rho_1)} \) with \( |\rho - \frac{\Re(\rho_1)}{\Im(\rho_1)}| < \epsilon_1 2^{-4} \) can be computed (the details are the same as in the proof of Lemma 5.4.6 where the ratio \( \sqrt[\rho_1]/\sqrt[\rho_2] \) has been approximated). Furthermore with Brent’s algorithm an approximation \( \phi_{\rho_1} \) to \( \arctan \rho \) with relative error less than \( \epsilon_1 2^{-6} \) is computed. Note that the images of \( \arctan \) are bounded in absolute value by \( \pi \). Therefore the absolute error between \( \phi_{\rho_1} \) and \( \arctan \rho \) is less than \( \epsilon_1 2^{-4} \).

Hence

\[
|\phi_\rho - \phi_{\rho_1}| \leq |\phi_\rho - \arctan \rho| + |\arctan \rho - \phi_{\rho_1}| < \epsilon_1 2^{-4} \leq \epsilon_1.
\]

For the bound on \( |\phi_\rho - \arctan \rho| \) we used the Mean-Value Theorem and the fact that the derivative of \( \arctan \) is bounded in absolute value by 1.

Then \( \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) is computed with absolute error less than \( \frac{1}{4} \epsilon_1 \). Since \( \phi_{\rho_1} < 4 \) an approximation to \( \frac{1}{\phi_{\rho_1}} \) with relative and hence absolute error less than \( \frac{1}{8} \epsilon_1 \) will suffice. Denote this approximation by \( \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \). Since \( |\frac{1}{\phi_{\rho_1}} - \frac{1}{\phi_{\rho_1}} \phi_{\rho_1}| \) the real number \( \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) is an approximation to \( \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) with absolute error less than \( \frac{1}{4} \epsilon_1 \).

Finally, we use Brent’s algorithm for sin and cos to compute \( \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) and sin \( \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) with relative and, since the sin and cos are bounded in absolute value by 1, absolute error less than \( \frac{1}{4} \epsilon_1 \). The derivatives of sin and cos are also bounded in absolute value by 1 so \( \cos \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) is an approximation to \( \cos \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) with absolute error less than \( \frac{1}{2} \epsilon_1 \). The same is true for the sin. Hence \( \cos \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} + i \sin \frac{1}{\phi_{\rho_1}} \phi_{\rho_1} \) has been approximated with absolute error less than \( \epsilon_1 \). The details of the analysis of this approximation algorithm for sin and cos are even simpler as the ones in the proof of Lemma 5.4.6, so we omit again the details.

Similarly we compute the approximation to \( \zeta_k \) from approximations to \( \pi \) and \( \zeta \). The analysis is as above.

As we mentioned above if these approximation are multiplied with the approximation for \( |\rho_1|^{\frac{1}{\zeta_k}} \) then the result is the desired approximation to \( \sqrt[\rho_1] \). This proves the first part of the lemma.

The second one can be shown in exactly the same way as the corresponding claim in Lemma 5.4.6.
If we compare the previous lemma with Lemma 5.4.6 we remark that complex radicals force us to use an initial approximation to \( \alpha \) that is roughly \( n \) times better than the one required for real radicals.

**Corollary 5.4.10** Approximations as required by Corollary 5.3.3 or by the complex case of Corollary 5.3.4 can be computed by \( \mathcal{O}(n) \) elementary operations on floating-point numbers of size \( \mathcal{O}(n^2 L) \), \( \mathcal{O}(\log n L) \) elementary operations on floating-point numbers of size \( \mathcal{O}(n L) \), and a constant number of operations on floating-point numbers of size \( \mathcal{O}(\log d + n L) \).

**Proof:** In Corollary 5.3.4 \( \epsilon \) is of order \( 2^{-\mathcal{O}(n L)} \). For Corollary 5.3.3 \( \epsilon \) is only of order \( 2^{-\mathcal{O}(n \bar{L})} \). But as follows from the proof of the previous lemma and the lower bound on the real part of an algebraic number as given in Lemma 5.4.8 we nevertheless have to work with approximations of order \( 2^{-\mathcal{O}(n L)} \) since otherwise we may be forced to divide by zero.

Let us finish this subsection by explicitly stating one result that has been shown while proving Lemma 5.4.6 and Lemma 5.4.9. It will be used on several occasion in the second part of this thesis.

**Lemma 5.4.11** Let \( \epsilon > 0 \). Suppose \( \rho \) is a complex number such that \( |\rho| < 2^{C_1} \) and \( |\Re(\rho)| > 2^{-C_2} \) and assume an approximation to \( \rho \) with absolute error less than \( \epsilon 2^{-(2C_1+2C_2+13)} \), is given. Then using \( \mathcal{O}(\log(\log \frac{1}{\epsilon} + C_1 + C_2)) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon} + C_1 + C_2) \) and \( \mathcal{O}(1) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon} + \log d) \) an approximation to \( \zeta_k \frac{|\rho|}{\tilde{\rho}} (\cos \frac{1}{d} \phi_\rho + i \sin \frac{1}{d} \phi_\rho), k < d, \) with absolute error less than \( \epsilon \) can be computed.

**Proof:** First of all exactly as in the beginning of the proof of Lemma 5.4.9 \( |\rho|^{\frac{1}{2}} \) is determined with absolute error less than \( \epsilon 2^{-(C_1+3)} \).

Next by the bounds on \((\Re(\rho))^{-1}\) and \(\Im(\rho)\) (the latter bound following from the bound on \(|\rho|\)) an approximation \( \rho \) to \( \frac{\Im(\rho)}{\Re(\rho)} \) with absolute error less than \( \epsilon 2^{-(C_1+7)} \) can be determined as follows.

If the approximation to \( \rho \) is denoted by \( \tilde{\rho} \) then \( |(\Re(\tilde{\rho}))^{-1} - (\Re(\rho))^{-1}| < \epsilon 2^{-(2C_1+10)} \) which follows from the fact that \( \Re(\rho) \) is approximated by \( \Re(\tilde{\rho}) \) with absolute error less than \( \epsilon 2^{-(2C_1+2C_2+13)} \) and Lemma 5.4.3. Hence approximating \( (\Re(\tilde{\rho}))^{-1} \) with absolute error less than \( \epsilon 2^{-(2C_1+10)} \) leads to an approximation of \( (\Re(\rho))^{-1} \) with absolute error less than \( \epsilon 2^{-(2C_1+9)} \).
Multiplying this approximation with $\Im(\tilde{\rho})$ yields the desired approximation $\rho$ to $\frac{\Im(\rho)}{|\Re(\rho)|}$. As is clear this approximation can be computed within the time bounds stated.

Then we compute the arctan of $\rho$ with absolute error less than $\epsilon 2^{-(C_1+7)}$. Denote the approximation by $\tilde{\phi}_\rho$.

As in the proof of Lemma 5.4.9 it can be shown that computing $\frac{1}{d}\tilde{\phi}_\rho$ with absolute error less than $\epsilon 2^{-(C_1+6)}$ gives an approximation to $\frac{1}{d}\phi_\rho$ with absolute error less than $\epsilon 2^{-(C_1+5)}$. As follows from Brent’s result these steps can also be done within the time bounds stated.

Again as in the proof of Lemma 5.4.9 Brent’s algorithm is used to compute $\cos\frac{1}{d}\phi_\rho + i \sin\frac{1}{d}\phi_\rho$ with absolute error less than $\epsilon 2^{-(C_1+3)}$.

Analogously, within the time bounds stated $\zeta^k_d$ is computed with absolute error less than $\epsilon 2^{-(C_1+3)}$.

Finally multiplying the approximations to $|\rho|^\frac{1}{d}$, $\cos\frac{1}{d}\phi_\rho + i \sin\frac{1}{d}\phi_\rho$, and $\zeta^k_d$ yields the approximation to $\zeta^k_d|\rho|^\frac{1}{d} (\cos\frac{1}{d}\phi_\rho + i \sin\frac{1}{d}\phi_\rho)$ with absolute error less than $\epsilon$. 

\[\square\]
5.5 A Probabilistic Test for Equality.

In the last two paragraphs we showed how to compute the coefficients of a number \( \gamma \in \mathbb{Q}(\alpha) \) such that if a radical \( \sqrt[p]{\gamma} \) or a ratio of radicals \( \sqrt[p_1]{\gamma}/\sqrt[p_2]{\gamma} \) is contained in \( \mathbb{Q}(\alpha) \) then \( \gamma = \sqrt[p]{\gamma} \) or \( \gamma = \sqrt[p_1]{\gamma}/\sqrt[p_2]{\gamma} \). But even if the algorithm outputs an element \( \gamma \) of \( \mathbb{Q}(\alpha) \) we cannot be sure whether these equalities are correct. In this paragraph it is shown how to use probabilistic methods to check in time polynomially in \( \log d_1, \log d_2 \) whether \( \gamma = \sqrt[p_1]{\gamma}/\sqrt[p_2]{\gamma} \) or \( \gamma = \sqrt[p]{\gamma} \). In the complex case this result is almost of no use since we assume in this case that \( \mathbb{Q}(\alpha) \) contains primitive \( d_1, d_2 \)-th roots of unity and therefore \( n = [\mathbb{Q}(\alpha) : \mathbb{Q}] \) may be of order \( \max\{d_1, d_2\} \) (see [Ap]). Therefore we also give a deterministic algorithm to decide whether \( \gamma = \sqrt[p_1]{\gamma}/\sqrt[p_2]{\gamma} \) or \( \gamma = \sqrt[p]{\gamma} \). This algorithm has run time which is polynomial only in \( d_1, d_2, \) or \( d \), respectively.

Observe that in order to check whether \( \gamma = \sqrt[p]{\gamma}/\sqrt[p_2]{\gamma} \), say, we only need to determine whether \( \gamma^{d_1 d_2} \rho_2^{d_1} = \rho_1^{d_2} \), because this implies that for some \( d_1 \)-th root of \( \rho_1 \) and some \( d_2 \)-th root of \( \rho_2 \) their ratio is in \( \mathbb{Q}(\alpha) \). Hence the ratio of all real roots of \( \rho_1, \rho_2 \) in the real case and the ratios of arbitrary roots of \( \rho_1, \rho_2 \) in the complex case are in \( \mathbb{Q}(\alpha) \). In particular, for the roots denoted by \( \sqrt[p_1]{\gamma} \) and \( \sqrt[p_2]{\gamma} \) the ratio \( \sqrt[p_1]{\gamma}/\sqrt[p_2]{\gamma} \) is in \( \mathbb{Q}(\alpha) \) and the algorithm of Corollary 5.3.4, p. 66, will return the correct representation for the ratio.

However, for Corollary 5.3.3, p. 66, the situation is different. If \( \gamma^d = \rho \) then \( \gamma \) need not be the \( d \)-th root \( \sqrt[p]{\gamma} \). Moreover, \( \mathbb{Q}(\alpha) \) need not contain all \( d \)-th roots of unity and it cannot be argued that if \( \mathbb{Q}(\alpha) \) contains one \( d \)-th root of \( \rho \) then it must contain all of them. Therefore even if \( \gamma^d = \rho \) we use a bit comparison test to decide whether \( \gamma = \sqrt[p]{\gamma} \) and hence if \( \sqrt[p]{\gamma} \in \mathbb{Q}(\alpha) \).

Let us begin by arguing why the usual methods to distinguish two algebraic numbers will not lead to an algorithm that is polynomial in the logarithm of the degrees \( d_1, d_2 \) of the radicals. First of all, in general we cannot bound the size of \( d_1, d_2 \) in terms of a polynomial expression in the input size of \( \mathbb{Q}(\alpha), \rho_1, \) and \( \rho_2 \). That is, we cannot give a polynomial upper bound such that if \( d_1, d_2 \) exceed this bound then the ratio \( \sqrt[p_1]{\gamma}/\sqrt[p_2]{\gamma} \) cannot be an element of \( \mathbb{Q}(\alpha) \). There are various methods to show that if \( d_1 \) or \( d_2 \) are larger than an expression polynomial in \( n, 2^l \), and \( 2^L \) then the ratio cannot be in \( \mathbb{Q}(\alpha) \). But this is clearly not good enough for our purposes.

Because there is no such upper bound on \( d_1, d_2 \) the usual root separation bounds to distinguish two algebraic integers will result in an algorithm that is polynomial only in \( d_1, d_2 \) or exponential in \( L, l \) rather than polynomial in
the input size. In fact, we have to distinguish the algebraic number \( \gamma \) from the ratio \( \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \) which may have degree \( d_1 d_2 \) over \( \mathbb{Q}(\alpha) \) so we may need at least \( nd_1 d_2 \) bits to separate these numbers.

We also cannot compute \( \rho_2^{d_1} \gamma^{d_1 d_2} \) by successive squaring and check whether it equals \( \rho_1^{d_2} \) because this would clearly imply that we had to work with numbers whose representations need \( \Omega(d_1 + d_2) \) bits.

The same argument applies already to simple radicals \( \sqrt[d]{\rho} \), in which case we had to check whether \( \gamma = \rho \). Although we know that the final result cannot have a large representation if the equation is correct, the intermediate results may require \( \Omega(d) \) bits. In particular, the denominator of the coefficients may get too large.

Instead of these approaches we describe an algorithm that actually uses successive squaring to check whether \( \gamma = \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \). But in order to avoid an exponential coefficient growth in the intermediate steps we reduce the coefficients modulo randomly chosen integers, that is, we use modular arithmetic. It will be shown that the error probability of the algorithm can be made exponentially small.

We claim that the following algorithm answers the question whether \( \gamma = \sqrt[d_1]{\rho_1}/\sqrt[d_2]{\rho_2} \) correctly with probability at least \( 1 - 2^{-t} \) for \( t \in \mathbb{N} \), \( t > 3 \).

Assume \( b_1, b_2, c \in \mathbb{Z} \) such that \( b_1 \rho_1 \in \mathbb{Z}[\alpha] \), \( c \gamma \in \mathbb{Z}[\alpha] \). Define \( \tilde{\rho}_i := b_i \rho_i \) and \( \tilde{\gamma} := c \gamma \). Then \( \sqrt[d_1]{\tilde{\rho}_1}/\sqrt[d_2]{\tilde{\rho}_2} = \gamma \) is equivalent to

\[
\tilde{b}_2 d_2 \tilde{\rho}_2^{d_1} \tilde{\gamma}^{d_1 d_2} - \tilde{b}_1 d_1 \tilde{\rho}_1^{d_1 d_2} = 0,
\]

which is an equation in \( \mathbb{Z}[\alpha] \). Denote the algebraic number on the left-hand side of the equation by \( \Gamma \in \mathbb{Z}[\alpha] \).

Fix an interval \( I = [1, 2^T] \), where \( T \) will be specified later. Randomly choose \( 6(t + 1)/T \) rational integers \( z_j \) from \( I \) and compute the coefficients of \( \Gamma \) modulo all integers \( z_j \). If for all \( j = 1, \ldots, 6(t + 1)/T \), the reduced coefficients of \( \Gamma \) are zero output \( \gamma = \sqrt[d_1]{\tilde{\rho}_1}/\sqrt[d_2]{\tilde{\rho}_2} \) otherwise output that the ratio of radicals \( \sqrt[d_1]{\tilde{\rho}_1}/\sqrt[d_2]{\tilde{\rho}_2} \) is not contained in \( \mathbb{Q}(\alpha) \).

To compute the coefficients of \( \Gamma \) modulo \( z_j \), \( j = 1, \ldots, 6(t + 1)/T \), first, using exact arithmetic in \( \mathbb{Z}[\alpha] \), the reduced coefficients of the numbers \( \tilde{b}_1 d_1, \tilde{b}_2 d_2, c d_1 d_2, \tilde{\rho}_1 d_1, \tilde{\rho}_2 d_2, \gamma d_1 d_2 \) are computed by successive squaring and reducing the coefficients of the intermediate results modulo \( z_j \). With these numbers the expression corresponding to \( \Gamma \) is formed and reduced, if necessary. The following lemma shows that this will give us the reduced coefficients of \( \Gamma \).
Lemma 5.5.1 Let \( \rho_1, \rho_2 \in \mathbb{Z}[\alpha] \), \( z \in \mathbb{Z} \), and denote by \((\rho_i)_z\) the number that is obtained by reducing the coefficients of \( \rho_i \) modulo \( z \). Then

\[
((\rho_1)_z + (\rho_2)_z)_z = (\rho_1 + \rho_2)_z
\]

and

\[
((\rho_1)_z (\rho_2)_z)_z = (\rho_1 \rho_2)_z
\]

using arithmetic in \( \mathbb{Z}[\alpha] \).

Proof: For the addition this is a straightforward computation.

For the multiplication we need to show that for two integer polynomials \( g_1, g_2 \)

\[
\left( \left( (g_1)_z \right)_z \left( (g_2)_p \right)_z \right)_z = \left( (g_1 g_2)_p \right)_z,
\]

where \((f)_p\) denotes the polynomial \( f \) reduced modulo the minimal polynomial \( p \) of \( \alpha \). Write \( g_i, i = 1, 2 \), as

\[
g_i = ph_i + r_i, \text{ deg } r_i < \text{deg } p,
\]

that is, \( r_i = (g_i)_p \).

Furthermore let

\[
r_i = r'_i + zr_i,
\]

with \( r'_i = (r_i)_z \).

Then

\[
\left( \left( (g_1)_z \right)_z \left( (g_2)_p \right)_z \right)_z = \left( (r'_1 r'_2)_p \right)_z.
\]

Obviously

\[
\left( (g_1 g_2)_p \right)_z = \left( (r_1 r_2)_p \right)_z.
\]

Now

\[
(r_1 r_2)_p = (r'_1 r'_2)_p + z(r'_1 r'_2 + r'_2 r'_1 + r'_1 r'_2)_p,
\]

since for a non-constant polynomial \( p \) the homomorphism \( \mathbb{Z}[X] \to (\mathbb{Z}[X])_p \) induces an isomorphism of \( \mathbb{Z} \) onto itself.

From the previous equality we finally get

\[
(r_1 r_2)_p = (r'_1 r'_2)_p,
\]

which proves the lemma. \( \square \)
Remark that
\[
\left( \left( \left( (g_1)_p \right)_z \left( (g_2)_p \right)_z \right)_p \right)_z = \left( (g_1 g_2)_p \right)_z
\]
is not correct if \( z \) is a non-constant polynomial.

It remains to analyze the run time and error probability of the algorithm. We begin with the run time.

We apply the following well-known result on the complexity of arithmetic in algebraic number fields (see [Lo1]).

**Lemma 5.5.2** Let \( \rho_1, \rho_2 \in \mathbb{Z}[\alpha] \) with \( [\rho_1], [\rho_2] < 2^B \). Furthermore assume that the minimal polynomial of the algebraic integer \( \alpha \) has degree \( n \) and satisfies \( |p| < 2^l \). Then

(i) the coefficients of \( \rho_1 + \rho_2 \) are bounded in absolute value by \( 2^{B+1} \) and \( \rho_1 + \rho_2 \) can be computed by \( n \) additions of integers of size \( B \),

(ii) the coefficients of \( \rho_1 \rho_2 \) are bounded in absolute value by \( 2^{2(nl+B)} \) and the product can be computed using \( O(n^2) \) elementary operations on integers of size \( O(nl + B) \).

The claim on the size of \( \rho_1 \rho_2 \) is a special case of a lemma that will be proven below.

Furthermore we need to reduce \( b, c \), the coefficients of \( \gamma' \), and the coefficients of the intermediate results by integers of size \( \leq 2^T \). Since by the previous lemma the coefficients of the intermediate results have bit size less than \( O(nl + T) \) each reduction step for a single coefficient can be done by a constant number of elementary operations on integers of size \( O(nl + T) \). Recall that reducing a coefficient modulo \( z_i \) is equivalent to a division with remainder. The overall number of these reduction steps is \( O(Tn(n \log d_1 + \log d_2)) \).

To analyze the initial reduction steps recall that if \( [\rho_i] < 2^L \) then \( [\gamma] < 2^{6L} \), \( L = [n \log n + nl + nL]^{12} \) (see Lemma 5.3.2, p. 64). Hence for a single \( z_i \) we can compute the initial reductions by \( O(n) \) elementary operations on integers of size \( O(L + T) \).

Finally, we need \( O(n^2(\log d_1 + \log d_2)) \) operations on integers of size \( O(nl + T) \) for the multiplications in \( \mathbb{Z}[\alpha] \).

We summarize this in

---

\[12\] If the numbers returned by the lattice reduction do not satisfy this bound then we may stop at once since the ratio of radicals will not be an element of \( \mathbb{Q}(\alpha) \).
Lemma 5.5.3 The algorithm described above uses $O(n(Tt + n)(\log d_1 + \log d_2))$ elementary operations on integers of size $O(L + T)$.

Next we analyze the error probability of the algorithm. First observe that if $\gamma = \sqrt[4]{\rho_1}/\sqrt[4]{\rho_2}$ then the algorithm above will always give the correct answer because all coefficients even if reduced modulo an integer will be zero. On the other hand, if the algorithm answers that $\gamma \neq \sqrt[4]{\rho_1}/\sqrt[4]{\rho_2}$ this answer is also correct since the algorithm found an integer such that the representations of $\gamma$ and the ratio of radicals differ already modulo this integer. So this number is a witness that $\gamma \neq \sqrt[4]{\rho_1}/\sqrt[4]{\rho_2}$.

Definition 5.5.4 If $\Gamma \neq 0$ then we call a number $z \in \mathbb{Z}$ unlucky if it divides all coefficients of $\Gamma$ or, equivalently, the gcd of these coefficients. Otherwise we call $z$ lucky.

Exactly the unlucky numbers will lead to an incorrect answer of the algorithm. Hence to prove the claim that the algorithm gives the correct answer with probability at least $1 - 2^{-t}$ we have to show that among $6(t + 1)$ randomly chosen integers from $I$ with probability at least $1 - 2^{-t}$ one number is lucky.

Obviously, we need a bound on the gcd of the coefficients in $\Gamma$.

Lemma 5.5.5 Let $\rho_1, \ldots, \rho_d \in \mathbb{Z}[\alpha]$ such that $|\rho_i| < 2^B$, $i = 1, 2, \ldots, d$, and $\alpha$ is as usual. Then the coefficients of $\prod_{i=1}^{d} \rho_i$ are bounded in absolute value by $2^{d(\log n + n(l+1)+B)}$.

Proof: Let $R_i(X) = \sum_{j=0}^{n-1} t_j^{(i)}X^j$ be defined by $R_i(\alpha) = \rho_i$. Hence computing the coefficients of $\prod_{i=1}^{d} \rho_i$ is the same as computing the coefficients of $\prod R_i(X) \mod p(X)$, where $p$ is the minimal polynomial of $\alpha$.

$\prod R_i(X)$ is a polynomial of degree $(n-1)d$ and its coefficients are bounded in absolute value by $2^{d(\log n+B)}$. Write $\prod R_i(X)$ as $\sum_{i=0}^{d(n-1)} m_i X^i$ and consider the following matrix:

$$
\begin{bmatrix}
1 & p_{n-1} & \cdots & p_1 & p_0 \\
1 & \cdots & p_1 & p_0 \\
\ddots & & \ddots & & \ddots \\
 & 1 & p_{n-1} & \cdots & p_1 & p_0 \\
 & m_{d(n-1)} & m_{d(n-1)-1} & \cdots & m_1 & m_0
\end{bmatrix}
$$

83
Using Gauss-elimination this matrix can be transformed into an upper triangular matrix

\[
\begin{bmatrix}
1 & p_{n-1} & p_{n-2} & \cdots & p_1 & p_0 \\
1 & p_{n-1} & \cdots & \cdots & p_1 & p_0 \\
\vdots & \vdots & \ddots & \cdots & \vdots & \vdots \\
0 & m'_{n-1} & m'_{n-2} & \cdots & m'_1 & m'_0 \\
\end{bmatrix}
\]

If follows that \(m'_i\) are the coefficients of \(\prod R_i(X)\) modulo \(p(X)\).

We have to analyze this process. Denote by \(B_i\) an upper bound on the absolute values on the entries in the last row after the \(i\)-th step of the Gauss-elimination. In particular, \(B_0 \leq 2^{d(\log n + B)}\). As is easily seen

\[
B_{i+1} \leq (|p|_\infty + 1)B_i.
\]

As we have to apply \(d(n-1) - n\) steps in the Gauss-elimination the lemma follows from \(|p|_\infty < 2^l\).

We apply this bound to \(b_1^{d_2} p_2^{d_1} c^{d_1 d_2} p_1^{d_2}\) and \(b_2^{d_1} c^{d_1 d_2} p_1^{d_2}\). Since \(|b_1|, |b_2|, |c|, |p'_1|, |p'_2|, |\gamma'| < 2^{6\mathcal{L}}\) (see Lemma 5.3.2, p. 64, and the remarks following this lemma) and \(d_1 + d_2 \leq d_1 d_2\) for \(d_1, d_2 \geq 2\), this shows that both expressions have coefficients whose absolute value is bounded by

\[
2^{2d_1 d_2 (\log n + n(l+1) + 6\mathcal{L})}.
\]

This finally yields a bound of

\[
C \leq 14d_1 d_2 \mathcal{L}
\]

for the bit size of the coefficients of \(\Gamma\). Hence the gcd of the coefficients in \(\Gamma\) is bounded by \(2^C\).

Unfortunately, an integer \(z \in \mathbb{Z}\) may have \(z^{\Omega(ln(\log z))}\) different divisors (cf. [Ch]). This is the main reason why we cannot show directly that most numbers in \(I\) are lucky. Instead we will show that most primes in \(I\) are lucky and that by choosing randomly \(6(t + 1)T\) numbers from \(I\) with high probability at least one of them is a lucky prime.
First note that any integer $z$ has at most $\log z$ different prime divisors. Hence the gcd of the coefficients of $\Gamma$ has at most $C$ distinct prime divisors. In general, this is a very crude estimate since we know that on the average the integer $z$ has only $\log \log z$ prime divisors (see [Ap]).

On the other hand there is the following well-known bound for the prime counting function $\pi(x)$. For a proof of this lemma see [Ap].

**Lemma 5.5.6** The number $\pi(2^T)$ of primes in the interval $[1, 2^T]$ is at least $\frac{1}{6\sqrt{T}}$.

This lemma has two consequences. It shows that if $T = 4(\log C + t)$ then the number of primes that do not divide the gcd of the coefficients of $\Gamma$ is at least $2^{\log C + t + 1}$. In fact, $T = 4(\log C + t)$ implies $T > \log C + \log T + t + 4$ and $\frac{1}{2T}/T > 2^{\log C + t + 1}$ for $t > 3$.

Hence at most a fraction of $2^{-t-1}$ of the primes in $I$ is unlucky. Equivalently,

**Lemma 5.5.7** Let $T = 4(\log C + t)$, $t > 3$. A randomly chosen prime in $I = [1, 2^T]$ is unlucky with probability less than $2^{-t-1}$.

As a second consequence to Lemma 5.5.6 we get the following lemma.

**Lemma 5.5.8** Let $I = [0, 2^T]$. If $6(t + 1)T$ numbers are chosen randomly from $I$ then with probability at least $1 - 2^{-(t+1)}$ one of the numbers is prime.

**Proof:** A random number in $I$ is composite with probability at most $1 - \frac{1}{T}$. Therefore the probability that none of the chosen numbers is prime is bounded by $(1 - \frac{1}{T})^{6(t+1)T}$.

\[
\left(1 - \frac{1}{6T}\right)^{6T} \leq \frac{1}{e},
\]

hence

\[
\left(1 - \frac{1}{6T}\right)^{6(t+1)T} \leq e^{-(t+1)} < 2^{-(t+1)}.
\]

The lemma follows.

Combining the two observations we have shown

**Lemma 5.5.9** Let $T = 4(\log(14d_1d_2\mathcal{L}) + t)$. If $6(t + 1)T$ integers are chosen randomly from $I = [1, 2^T]$ then with probability at least $1 - 2^{-t}$ one of the integers is lucky.
Proof: There are two ways we may fail to hit upon a lucky integer. First no prime may have been chosen. Second, even if a prime has been chosen it may not be lucky. Both cases happen independently and with probability at most $2^{-(t+1)}$.

Finally we combine Lemma 5.5.3 with the previous lemma and state the run time for a deterministic test. The analysis of the deterministic test is straightforward. We check whether $\gamma = \frac{d_1}{d_2} \sqrt{\rho_1}/\sqrt{\rho_2}$ but raising both sides of the equation to the lcm$(d_1, d_2)$-th power and apply Lemma 5.5.2. For the probabilistic algorithm we may also have taken lcm$(d_1, d_2)$-th powers but in that case asymptotically it saves us nothing.

Corollary 5.5.10 Given the element $\gamma = \frac{1}{n} \sum_{i=0}^{n-1} c_i \alpha^i$ from Corollary 5.3.4 it can be checked with error probability less than $2^{-t}$ and using at most $O(n \max\{\log d_i\}(n + t(\log \mathcal{L} + \max\{\log d_i\})))$ elementary operations on integers of size $O(\mathcal{L} + \max\{\log d_i\} + t)$ whether $\gamma = \frac{d_1}{d_2} \sqrt{\rho_1}/\sqrt{\rho_2}$. Furthermore, with $O(\max\{\log d_i\} n^2)$ elementary operations on integers of size $O(\text{lcm}(d_1, d_2) \mathcal{L})$ it can be checked deterministically whether the element $\gamma$ of Corollary 5.3.4 satisfies $\gamma = \frac{d_1}{d_2} \sqrt{\rho_1}/\sqrt{\rho_2}$.

Note that the algorithm above resembles in some respects Brown’s modular gcd-algorithm [Bro]. But in the brute force non-modular gcd algorithm only the intermediate results may have exponential size while the gcd itself has polynomial size. In our case, both intermediate and final results, if not reduced, may have exponential size. This is the main reason why the algorithm above is probabilistic and Brown’s modular gcd-algorithm is not.

As mentioned in the complex case the probabilistic algorithm is almost of no use.

Observe that of course for testing whether simple radicals $\sqrt[d]{\rho} \in \mathbb{Q}(\alpha)$ similar results apply. Actually by choosing $d_2 = 1$ they are covered by the previous theorems. But in Corollary 5.3.3, p. 66, we did not require that $\mathbb{Q}(\alpha)$ contains a primitive $d$-th root of unity. Hence even if $\gamma^d = \rho$ it need not be the case $\sqrt[d]{\rho} = \gamma$ for the specific $d$-th root $\sqrt[d]{\rho}$ of the corollary.

But $\gamma$ will be a $d$-th root of $\rho$. Therefore $|\gamma - \sqrt[d]{\rho}| = |\zeta_d - 1| \sqrt[d]{\rho}$. Observe that two $d$-th roots of unity have distance at least $2^{-1-\log d+\log \pi}$, which is easily seen by recalling that the $d$-th roots of unity correspond to the vertices of a regular $d$-gon inscribed in the unit circle in the plane. Hence computing the difference $\gamma - \sqrt[d]{\rho}$ with error less than $2^{-\mathcal{L}-\log d}$ will show whether $\gamma = \sqrt[d]{\rho}$. This test is necessary only if $\log d$ is larger than $n, l$, and
since otherwise the approximation required by the reconstruction step guarantees already $\gamma = \sqrt[d]{\rho}$ if $\gamma^d = \rho$.

**Corollary 5.5.11** Given an approximation with absolute error less than $2^{-L-\log d}$ to the element $\gamma$ from Corollary 5.3.3 then it can be decided with error probability less than $2^{-t}$ and using at most $O(n \log d(n + t(t + \log L + \log d)))$ elementary operations on integers of size $O(L + \log d + t)$ whether $\gamma = \sqrt[d]{\rho}$. Furthermore, with $O(n^2 \log d)$ elementary operations on integers of size $O(dL)$ it can be checked deterministically whether the element $\gamma$ of Corollary 5.3.3 satisfies $\gamma = \sqrt[d]{\rho}$.

Finally, let us mention that for algebraic number fields whose ring of integers is a unique factorization domain and the (then well-defined) gcd of two integers can be computed efficiently then the test can be made deterministic in any case. In fact, then the same algorithm as for the rational numbers can be used. It can be shown that only the exponential growth in the denominator of an algebraic number when raised to a $d$-th power forced us to use probabilistic methods. But for algebraic integers the denominator is always bounded by the discriminant of the number field.
5.6 Sums of Radicals over Algebraic Number Fields

We summarize the results of the previous subsections in the following table.

Table 1: Run Times for Ratios of Radicals

<table>
<thead>
<tr>
<th>procedure</th>
<th>number of operations</th>
<th>bit size of numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximations to $\sqrt[d_i]{\rho_i}$</td>
<td>$O(n)$</td>
<td>$O(n^2L)$</td>
</tr>
<tr>
<td></td>
<td>$O(1)$</td>
<td>$O(nL + \max{\log d_i})$</td>
</tr>
<tr>
<td></td>
<td>$O(\log nL)$</td>
<td>$O(nL)$</td>
</tr>
<tr>
<td>Reconstruction</td>
<td>$O(n^3L)$</td>
<td>$O(nL)$</td>
</tr>
<tr>
<td>Probabilistic test ($&lt; 2^{-t}$)</td>
<td>$O(n \max{\log d_i})$</td>
<td>$O(t + L + \max{\log d_i})$</td>
</tr>
<tr>
<td></td>
<td>$(n + t(t + \log L + \max{\log d_i}))$</td>
<td></td>
</tr>
<tr>
<td>Deterministic test</td>
<td>$O(\max{\log d_i}n^2)$</td>
<td>$O(\lcm(d_1d_2)L)$</td>
</tr>
</tbody>
</table>

Only for the approximation algorithms floating-point numbers are required.

Theorem 5.6.1 Let $\mathbb{Q}(\alpha)$ be a real algebraic number field, where $\alpha$ is an algebraic integer whose minimal polynomial is of degree $n$ and has length $|p|_2 < 2^t$. Furthermore let $\{\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}\}$ be a set of positive real radicals over $\mathbb{Q}(\alpha)$, where $|\rho_i| < 2^L$, $i = 1, 2, \ldots, k$.

It can be decided with error probability less than $2^{-t}$ and by using at most $O(k^2n^3L + k^2n \max\{\log d_i\}(n + t(\max\{\log d_i\} + \log L + t)))$ elementary operations on floating-point numbers or integers of size $O(nL + t + \max\{\log d_i\})$ whether the set of radicals is linearly independent. Within the same error and time bounds a maximal linearly independent subset of radicals can be computed. Here $L = \lceil n \log n + nL \rceil$.

Furthermore, given a linear combination $S = \sum_{i=1}^{k} v_i \sqrt[d_i]{\rho_i}$ of the radicals over $\mathbb{Q}(\alpha)$, $v_i \in \mathbb{Q}(\alpha)$, $|v_i| < 2^L$, then it can be decided with the same error probability and using additional $O(kn)$ elementary operations on integers of size $O(kL)$ whether $S = 0$.

Proof: To prove the first claim we apply the algorithms of the previous subsections to the $\frac{k(k-1)}{2}$ different pairs of radicals in $\{\sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k}\}$.
Furthermore, in this case for each pair we choose the error bound to be $2^{-t-2\log k}$. Hence the overall error probability is bounded by $2^{-t}$. Also note that the approximation algorithms to $\alpha$ and the radicals $\sqrt[p_i]{\rho_i}$ need to be applied only once.

To prove the second claim use this algorithm to partition \{ $\sqrt[p_1]{\rho_1}$, $\sqrt[p_2]{\rho_2}$, ..., $\sqrt[p_k]{\rho_k}$ \} into subsets $R_1, \ldots, R_h$ such that two radicals are in the same subset if and only if their ratio is an element of $\mathbb{Q}(\alpha)$. To simplify the notation assume $\sqrt[p_i]{\rho_i} \in R_i$, $i = 1, \ldots, h$. By the above result this partitioning can be done within the time bounds stated. Also elements $\nu_{ij}$ can be computed such that if $\sqrt[p_j]/\sqrt[p_i] \in \mathbb{Q}(\alpha)$ then $\sqrt[p_j]/\sqrt[p_i] = \nu_{ij}$. So

$$S = \sum_{i=1}^{k} v_i \sqrt[p_i]{\rho_i} = \sum_{i=1}^{h} \left( \sum_{\sqrt[p_j]{\rho_j} \in R_i} v_j \nu_{ij} \right) \sqrt[p_i]{\rho_i}.$$  

Since for any pair of radicals in $R' = \{ \sqrt[p_1]{\rho_1}, \sqrt[p_2]{\rho_2}, \ldots, \sqrt[p_k]{\rho_k} \}$ their ratio is not in $\mathbb{Q}(\alpha)$ we conclude by Corollary 3.10, p.28, that $S = 0$ if and only if

$$\sum_{\sqrt[p_j]{\rho_j} \in R_i} v_j \nu_{ij} = 0 \text{ for } i = 1, \ldots, h.$$  

It remains to show that the exact representations of these sums as elements in $\mathbb{Q}(\alpha)$ can be computed within the time bounds stated in the theorem. With these representations it is trivial to check whether the sums are zero.

By assumption $[v_j] < 2^L$ and by Lemma 5.3.2, p. 64, $[\nu_{ij}] < 2^{O(L)}$. Hence by Lemma 5.5.2, p. 82, each product $v_j \nu_{ij}$ can be computed by $O(n^2)$ elementary operations on integers of length at most $O(L)$, furthermore, the representation size of the product is bounded by $O(L)$. Observe that at most $k - 1$ products have to be computed. Hence for this step $O(kL)$ elementary operations on integers of size $O(L)$ are needed which is already covered by the previous run time bound.

Next we compute for all $i = 1, \ldots, h$, the products of the denominators of the products $v_j \nu_{ij}$. These denominators are bounded in absolute value by $O(L)$. Hence for all sums the corresponding products can be computed by $O(k)$ elementary operations on integers of size $O(kL)$. With these products it is easy to compute the representations of the sums as linear combinations of the basis elements $\alpha^i$ by $O(kn)$ elementary operations on integers of size $O(kL)$ (see Lemma 5.5.2, p. 82). This yields the additional operations mentioned in the theorem.
The statement on the error probability follows directly from the error probability for determining a maximal linearly independent subset.

The important thing about this theorem is that the run time is polynomial in the input size of the problems.

If we want the algorithm only to be polynomial in the \( d_i \)'s the algorithm can easily be made deterministic. Instead of using the probabilistic algorithm of the previous subsection we use the deterministic one.

\textbf{Theorem 5.6.2} If the tests from the theorem above are made deterministic then the algorithm uses at most \( O(k^2n^3L + k^2n^2 \max \{\log d_i\}) \) elementary operations on floating-point numbers and integers of size at most \( O((n + \max \{d_i^2\} + k)L) \).

Observe that in the last two theorems the dependence on the size of the coefficients \( \upsilon_i \) in the sum \( \sum_{i=1}^k \upsilon_i \sqrt[d_i]{\rho_i} \) is better than the dependence on the size of the elements \( \rho_i \). Even if these coefficients have representation size \( O(nL) \) the upper bounds given on the number of bit operations still apply.

Finally we let us state similar but more restricted results for complex radicals over number fields containing appropriate roots of unity. The proof is exactly as above.

\textbf{Theorem 5.6.3} Let \( \mathbb{Q}(\alpha) \) be an algebraic number field containing primitive \( d_i \)-th roots of unity \( i = 1, 2, \ldots, k \), where \( \alpha \) is an algebraic integer whose minimal polynomial is of degree \( n \) and has length \( |p|_2 < 2^l \). Furthermore let \( \{ \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_k]{\rho_k} \} \) be a set of radicals over \( \mathbb{Q}(\alpha) \), where \( |\rho_i| < 2^L \), \( i = 1, 2, \ldots, k \).

Using at most \( O(k^2n^3L + k^2n^2 \max \{d_i\}) \) elementary operations on floating-point numbers and integers of size \( O((n + \max \{d_i^2\})L) \) it can be decided whether the set of radicals is linearly independent over \( \mathbb{Q}(\alpha) \). Within the same error and time bounds a maximal linearly independent subset of radicals can be computed. Here \( L = [n \log n + nl + nL] \).

Furthermore, given a linear combination \( S = \sum_{i=1}^k \upsilon_i \sqrt[d_i]{\rho_i} \) of the radicals over \( \mathbb{Q}(\alpha) \), \( \upsilon_i \in \mathbb{Q}(\alpha) \), \( |\upsilon_i| < 2^L \), then it can be decided using additional \( O(kn) \) elementary operations on integers of size \( O(kL) \) whether \( S = 0 \).

Let us note that by assuming that \( \mathbb{Q}(\alpha) \) contains \( d_i \)-th roots of unity we may have to work in a field whose degree is exponential in \( k \). But for a small number of different degrees this may still be efficient. In particular, this is
true if all \( d_i \) are equal. The following example is one important application of this case.

Assume we are given a sum of complex \( d \)-th roots of rational numbers and want to decide whether it is zero. We can solve this problem by applying Theorem 5.6.3 with \( \mathbb{Q}(\alpha) \) being the \( d \)-th cyclotomic field. This allows us to decide the question in time polynomial in \( d, k \). The brute force approach (computing enough bits) has run time \( \Omega(d^k) \). Hence we have the following partial generalization of Corollary 4.1.5, p.39.

**Corollary 5.6.4** Assume \( S = \sum_{i=1}^{k} v_i \sqrt[d]{q_i} \) is a sum of radicals over \( \mathbb{Q} \) such that \( v_i, q_i \) are rational numbers whose numerator and denominator are bounded in absolute value by \( 2^L \). Then it can be decided using \( \mathcal{O}(k^2d^4(d + L)) \) elementary operations on floating-point numbers and integers of size \( \mathcal{O}(d^3 + d^2L) \) and \( \mathcal{O}(kd) \) elementary operations on integers of size \( \mathcal{O}(kd^2 + kdL) \) whether \( S = 0 \).

**Proof:** Applying Mignotte’s bounds [Mi1] for the size of the coefficients of a factor of a polynomial to \( X^d - 1 \) and the irreducible polynomial of a primitive \( d \)-th root of unity shows that the latter has length bounded by \( 2^{d+1} \). Since this polynomial has degree \( \phi(d) < d \) the corollary follows from the previous theorem by observing that in this case the deterministic test is applied only to radicals of equal degree. Hence \( \max\{d_i^2\} \) can be replaced by \( d \) (see Table 1).

If \( \mathbb{Q}(\alpha) \) contains a primitive \( d \)-th root of unity then applying Theorem 5.6.3 to the set \{1, \( \sqrt[d]{\rho} \)\} simply checks whether \( \sqrt[d]{\rho} \in \mathbb{Q}(\alpha) \). In the denesting algorithms to be described in the next sections exactly this situation will occur.

However, in Corollary 5.3.3, p. 66, we did not assume that \( \mathbb{Q}(\alpha) \) contains primitive \( d \)-th roots of unity. By Corollary 5.5.11 in this case we need to approximate the candidate \( \gamma \) and \( \sqrt[d]{\rho} \) with absolute error less than \( 2^{-L-\log d} \).

Since \( |\gamma| < 2^{dL} \) (Lemma 5.3.1, p. 63) changing in Table 2 the last entry for the approximation algorithm to “\( \mathcal{O}(\log nL) \) operations on floating-point numbers of size \( \mathcal{O}(nL + \log d) \)” takes care of this. This proves the next theorem. It states for example how much time it takes to check whether

\[13\]For the reconstruction and verifying step \( L \) can also be replaced by \( \tilde{L} \) (see Lemma 5.3.1, p. 63, and Corollary 5.3.3, p. 66).
a field contains certain primitive roots of unity. Therefore whenever we assume that a number field contains some root of unity this can be checked efficiently.

**Theorem 5.6.5** Let \( Q(\alpha) \) be an algebraic number field, where \( \alpha \) is an algebraic integer with minimal polynomial \( p \) of degree \( n \) and length bounded by \( 2^L \). Let \( \rho \in Q(\alpha) \) satisfy \( |\rho| < 2^L \). For \( d \in \mathbb{N} \) it can be decided with error probability less than \( 2^{-t} \) and using at most \( O(n^3L + n \log d(n + t(\log d + \log L + t))) \) elementary operations on floating-point numbers and integers of size \( O(nL + t + \log d) \) whether \( \sqrt[d]{\rho} \in Q(\alpha) \).

The deterministic test whether \( \sqrt[d]{\rho} \in Q(\alpha) \) uses at most \( O(n^3L + n^2d) \) elementary operations on floating-point numbers and integers of size \( O((n + d)L) \).
6 Denesting Radicals - The Basic Results

In the second part of this thesis we consider a problem which is known as *denesting of nested radicals* and has attracted a lot of mathematicians and computer scientists throughout the last years (see [BFHT],[HH],[La2],[La3],[Z]). Before stating the problem formally let us give some examples found in the notebook of the Indian mathematician Ramanujan [R].

\[
\sqrt[3]{\sqrt{28} - 3} = \frac{1}{3} \left( \sqrt[3]{98} - \sqrt[3]{28} - 1 \right)
\]

\[
\sqrt[5]{\sqrt[3]{5} \sqrt{32} - 5} = \sqrt[5]{1/25 + \sqrt{3/25} - \sqrt{9/25}}.
\]

In both equations the formula on the left-hand side has nesting depth 2 and the formula on the right-hand side has depth 1.

Although these examples may already explain the notion of nesting depth sufficiently we give a more formal definition (see also [BFHT]).

**Definition 6.1** The nesting depth of an expression over a field \(F\) is defined as follows:

1. an element of \(F\) has nesting depth 0 over \(F\),
2. an arithmetic combination \((A + B, A - B, A \times B, A/B)\) of expressions \(A, B\) over \(F\) has nesting depth \(\max\{\text{depth}(A), \text{depth}(B)\}\), and finally,
3. a root \(\sqrt[d]{A}\) of an expression \(A\) has nesting depth \(\text{depth}(A) + 1\).

Of course, given a nested radical there is no unique number in \(\mathbb{C}\), say, corresponding to this expression since the roots can be interpreted in different ways. Instead it always has to be said which roots are meant so that the value \(v(A)\) of a nested radical is uniquely defined. Then the problem of denesting nested radicals over a field \(F\) can be defined in the following way. Given a radical expression \(A\) over a field \(F\) with uniquely defined value \(v(A)\). Is there an expression \(B\) over \(F\) with the same value as \(A\) and with lower nesting depth over \(F\) than \(A\)?

Borodin et al. [BFHT] considered certain depth 2 expressions involving real square roots and showed under which conditions such an expression can be denested. To be more specific they proved the following theorem.

**Theorem 6.2 (Borodin et al.)** Let \(F \supset \mathbb{Q}\) be a real field. Let \(\alpha, \beta, \rho\) in \(F\) with \(\sqrt{\rho}\) not in \(F\). If \(\sqrt[3]{\alpha + \beta \sqrt{\rho}}\) is contained in some real radical extension
$ F( \sqrt[d]{\rho_1}, \ldots, \sqrt[m]{\rho_m}, \rho_1, \ldots, \rho_m \in F, \text{ of } F \text{ then for some positive } \gamma_0 \in F$

either

$$\sqrt{\gamma_0} \sqrt[\alpha + \beta \sqrt{\rho}]{} \in F(\sqrt{\rho})$$

or

$$\sqrt{\gamma_0} \sqrt{\rho} \sqrt[\alpha + \beta \sqrt{\rho}]{} \in F(\sqrt{\rho})$$

In the first case $\sqrt[\alpha^2 - \beta^2 \rho]{} \in F$ and $\gamma_0$ can be chosen as $2(a + \sqrt{\alpha^2 - \beta^2 \rho})$. In the second case $\sqrt[\rho \left(\beta^2 \rho - \alpha^2\right)]{} \in F$ and $\gamma_0$ can be chosen as $2(b + \sqrt{\rho \left(\beta^2 \rho - \alpha^2\right)})$.

Clearly, this theorem almost immediately leads to an algorithm computing a denesting. Since $\gamma_0$ is known and the field has the basis $\{1, \sqrt{\rho}\}$ it remains to determine the corresponding element in $F(\sqrt{\rho})$, which is not too difficult if $F$ is for example $\mathbb{Q}$.

This procedure always finds a denesting using only sums of depth 1 radicals. But observe that if a radical can be denested by a rational expression in radicals of smaller nesting depth then it can be denested by a sum of radicals of smaller nesting depth. In fact, simply consider the field $E$ containing all radicals appearing in the expression denesting the original radical expression $\gamma$. $\gamma$ is an element of $E$ which has a basis of products of the radicals of smaller depth than the depth of $\gamma$.

A few years after Borodin et al. published their results Landau [La2] showed how to compute a minimum depth denesting that is just one off the optimum and Horng, Huang [HH] showed how to compute the minimum depth denesting of an arbitrary nested radical over a field containing all roots of unity. Unfortunately, they were only able to prove single-exponential or double-exponential bounds on the output size. As the input size they consider the size of the minimal polynomial of the expression that has to be denested. This size may already be exponential in the number of bits necessary to describe the radical expression itself.

In particular, Horng and Huang showed that the minimum denesting of a nested radical defined over a field containing all roots of unity can already be found using only elements of a field $F$ that is generated by a root of unity whose degree over $\mathbb{Q}$ is in the worst case double-exponential in the input size of the minimal polynomial of the nested radical. Accordingly, the algorithms, which are completely due to Landau, have a double-exponential run time. Basically these algorithms work by first computing the Galois group of the minimal polynomial and then applying the classical results on solvable groups.
Although quite general these results cannot explain or determine the
denestings of Ramanujan mentioned in the beginning simply because these
examples use no roots of unity at all. Moreover, a double-exponential algo-
rithm is of almost no practical use.

By going back to the work of Borodin et al. [BFHT], simplifying, and gen-
eralizing their results, in this thesis we completely determine the structure
of denestings as the ones found by Ramanujan. In particular, we generalize
the theorem mentioned above to arbitrary depth 2 expressions containing
real radicals. These radicals need not be square roots.

Based on these results we also achieve efficient denesting algorithms for
a large class of depth 2 expressions. In particular, the run times will be at
most polynomial and in many cases even less than polynomial in the input
size of the minimal polynomial of the radical expression to be denested. The
class of depth 2 radicals to which these algorithms can be applied contains all
elements given by Ramanujan and is not restricted to expressions involving
only real radicals. Since we describe the algorithms for algebraic number
fields they can be applied repeatedly to nested radicals of depth larger than
2 in order to find non-trivial denestings for arbitrary radical expressions
efficiently. Although in general the algorithms cannot find a minimum depth
denesting in many cases they will do.

The denesting algorithms can also be used to generalize the results of
Section 5 in the following sense.

Suppose we want to apply the results of the previous section to the sum

$$S = \sum_{i=1}^{k} \sqrt[4]{a_i} + b_i \sqrt[8]{q_i}, a_i, b_i, q_i \in \mathbb{Q}.$$ 

We always assumed that the expressions under the root signs are elements of
a single extension $\mathbb{Q}(\alpha)$ of $\mathbb{Q}$. Hence in case of the sum $S$ we had to consider
the field generated by the radicals $\sqrt[4]{q_1}, \sqrt[8]{q_2}, \ldots, \sqrt[2^{2k}]{q_k}$, whose degree
will in general be exponential in $k$ even if the $d_{ij}$’s are constant. Therefore
the run times of the algorithms will be exponential in $k$.

Using the results of the following sections we can solve this problem
much more efficiently. First we check whether the sum $S$ can be written
as a sum of radicals $\sqrt[n]{p}$ over $\mathbb{Q}$. If this is not the case, $S$ cannot be zero.
On the other hand, if $S$ is transformed into a sum $\sum c_i \sqrt[n]{p}$ of radicals over
$\mathbb{Q}$ determining whether it is zero can easily be done by the algorithms of
Section 5. Moreover, as the size of $\sum c_i \sqrt[n]{p}$ will be small the algorithm will
do so efficiently.
Hence all we have to do is to determine upper bounds on the representation size of \( \sum c_i \sqrt[p_i]{\rho_i} \) and to describe an algorithm that computes it.

The following two sections of this thesis are organized as follows. In the first part of the next section we prove certain generalizations of the basic theorem of [BFHT]. In the second part we provide the basic means for our denesting algorithms and give a first description of these algorithms. They will be described in more detail in Section 7 where we also give an analysis of the algorithms. This part is quite technical and hopefully the reader will be already convinced at the end of Section 6 that our algorithms are much more efficient than the general algorithms known so far.

6.1 The Basic Theorems

In this section we prove our basic theorems on denesting nested radical expressions. For the time being we denote by a symbol \( \sqrt[p_i]{\rho_i}, d \in \mathbb{N}, \rho \in \mathbb{C} \), a solution to the equation \( X^d - \rho = 0 \).

If we do not restrict the radical then \( \sqrt[p_i]{\rho_i} \) denotes any of the \( d \) different solutions of \( X^d - \rho = 0 \). If we require that the radical \( \sqrt[p_i]{\rho_i} \) is real then \( \sqrt[p_i]{\rho_i} \) denotes one of the at most two real solutions of \( X^d - \rho = 0 \).

Hence it is implicitly assumed that \( \rho \in \mathbb{R} \) and that \( X^d - \rho = 0 \) has a real solution.

First we restrict ourselves to a single nested radical.

Given a nested radical, like for example \( \sqrt[p_{i_1}]{\sum_{k=1}^{k} \kappa_i \sqrt[p_i]{\rho_i}} \), of depth 2 over some field \( F \). To describe and compute the possible denestings of this nested radical we consider \( \gamma = \sum_{k=1}^{k} \kappa_i \sqrt[p_i]{\rho_i} \) as an element of the radical extension \( E = F(\sqrt[p_{i_1}]{\rho_1}, \ldots, \sqrt[p_{i_k}]{\rho_k}) \) generated by the depth one radicals \( \sqrt[p_i]{\rho_i} \) appearing in \( \sum_{k=1}^{k} \kappa_i \sqrt[p_i]{\rho_i} \). The theorems that we are going to prove in this section show that this field is almost the right place to look for denestings.

Observe that if \( F \subset \mathbb{R} \), \( \sqrt[p_i]{\rho_i} \in \mathbb{R}, i = 1, 2, \ldots, k \), and the radicals \( \sqrt[p_i]{\rho_i} \) are linearly independent over \( F \) then \( E = F(\gamma) \) (Theorem 3.13, p. 32). The same is true for not necessarily real radicals \( \sqrt[p_i]{\rho_i} \) if \( F \) contains primitive \( d_i \)-th roots of unity (Theorem 3.11, p. 31).

Let us fix some notation. Throughout this subsection \( F \) is a subfield of \( \mathbb{C} \) and \( E = F(\sqrt[p_{i_1}]{\rho_1}, \ldots, \sqrt[p_{i_k}]{\rho_k}) \) is a radical extension of \( F \). By \( N = [E : F] \) denote the degree of \( E \) over \( F \). By \( n_j \) denote the degree of the extension \( F(\sqrt[p_{i_1}]{\rho_1}, \ldots, \sqrt[p_{i_j}]{\rho_j}): F(\sqrt[p_{i_1}]{\rho_1}, \ldots, \sqrt[p_{i_{j-1}}]{\rho_{j-1}}) \). As has been mentioned in Section 2 a basis for the extension \( E : F \) is given by

\[
B = \{\beta_0, \beta_1, \ldots, \beta_{N-1}\} = \left\{ \prod_{j=1}^{k} \sqrt[p_j]{\rho_j^{e_j}}, \ 0 \leq e_1 < n_1, \ldots, 0 \leq e_k < n_k \right\}.
\]
As before $B$ is called the standard basis of $E : F$.

The next theorem generalizes and simplifies the proof of the result due to Borodin et al. [BFHT] (Theorem 6.2).

**Theorem 6.1.1** Let $F$ be a subfield of $R$, $d_i \in N$, $\rho_i \in F$, $\sqrt[di]{\rho_i} \in R$, $\sqrt[di]{\rho_i} > 0$ for all $i = 1, 2, \ldots, k$, $\gamma \in E = F(\sqrt[di]{\rho_i}, \ldots, \sqrt[di]{\rho_k})$, $d \in N$.

Assume $\sqrt[d]{\gamma}$ is real and denests over $F$ using only real radicals, that is, $\sqrt[d]{\gamma} \in F(\sqrt[ti]{\gamma_1}, \ldots, \sqrt[tm]{\gamma_m})$, for some $t_i \in N$, $\gamma_i \in F$, $p_i > 0$. Then there exist $\gamma_0 \in F \setminus \{0\}$, $\beta_j \in B$ such that

$$\sqrt[d]{\gamma_0} \sqrt[d]{\beta_j} \sqrt[d]{\gamma} = \eta \in E,$$

where the real $d$-th roots of $\gamma_0$, $\beta_j$ are implied in the above expressions.

Hence writing $\eta$ as a linear combination of the elements of the standard basis $B$ of $E$ and dividing by $\sqrt[d]{\gamma_0} \sqrt[d]{\beta_j}$ leads to a denesting of $\sqrt[d]{\gamma}$.

Observe that the condition $\sqrt[di]{\rho_i} > 0$ is no restriction since any real radical extension can clearly be generated by positive radicals. However, if it is generated by positive radicals then any element of the standard basis is a positive real number. Hence the real $d$-th root of $\beta_j$ as required by the theorem will always exist.

**Proof:** $\sqrt[d]{\gamma} \in F(\sqrt[ti]{\gamma_1}, \ldots, \sqrt[tm]{\gamma_m})$ clearly implies $\sqrt[d]{\gamma} \in E(\sqrt[ti]{\gamma_1}, \ldots, \sqrt[tm]{\gamma_m})$. Hence $\sqrt[d]{\gamma}$ is a radical over $E$ contained in $E(\sqrt[ti]{\gamma_1}, \ldots, \sqrt[tm]{\gamma_m})$. Since both fields, $E$ and $E(\sqrt[ti]{\gamma_1}, \ldots, \sqrt[tm]{\gamma_m})$, are real we can apply Lemma 3.6, p. 26, which shows

$$\sqrt[d]{\gamma} = \eta \prod_{i=1}^{m} \sqrt[d]{\gamma_i}^{f_i}, \text{ or } \sqrt[d]{\gamma} \left( \prod_{i=1}^{m} \sqrt[d]{\gamma_i}^{f_i} \right)^{-1} = \eta$$

for some $\eta \in E$, $f_i \in N$.

$\prod_{i=1}^{m} \sqrt[d]{\gamma_i}^{f_i}$ can be written as $\sqrt[d]{\rho}$ for some $\rho \in F$ and $t = \prod_{i=1}^{m} t_i$. Taking $d$-th powers this yields

$$\gamma/\eta^d = \sqrt[d]{\rho^d}.$$

But $\gamma/\eta^d \in E \subseteq R$ and $\rho \in F$. Therefore $\sqrt[d]{\rho^d}$ is a real radical over $F$ that is contained in $E \subseteq R$. By Lemma 3.6, p. 26, $\sqrt[d]{\rho^d} = \gamma_0/\beta_h$ for $\gamma_0 \in F$ and some basis element $\beta_h$.

Since $\sqrt[d]{\rho} = \frac{1}{\eta} \sqrt[d]{\gamma} \in R$ this shows that $\sqrt[d]{\rho}$ can be written as a real $d$-th root of $\gamma_0/\beta_h$. By assumption $\beta_h$ is positive and therefore $\gamma_0$ must be positive.

97
if $d$ is even. Hence real $d$-th roots of $\gamma'_0$ and $\beta_h$ exist such that

$$\sqrt[1/d]{\frac{1}{\gamma'_0}} \sqrt[1/d]{\frac{1}{\beta_h}} \sqrt[1/d]{\gamma} = \eta \in E.$$ 

Choosing different real roots of $\gamma'_0$ and $\beta_h$ just effects the sign. Therefore multiplying $\sqrt[1/d]{\gamma}$ with arbitrary real $d$-th roots of $\gamma'_0, \beta_h$ leads to an element of $E$.

Finally observe that $\beta_{h-1}$ is a radical of $E$ over $F$. Hence it can be written as $\sqrt[1/d]{\gamma'_0} \sqrt[1/d]{\beta_j}$ for a basis element $\beta_j \in B$ and $\gamma'_0 \in F$, $\gamma'_0 > 0$. So $\gamma_0$ may be chosen as $\sqrt[1/d]{\gamma'_0} \neq 0$ to prove the first claim.

To prove that the equation $\sqrt[1/d]{\gamma_0} \sqrt[1/d]{\beta_j} \sqrt[1/d]{\gamma}$ leads to a denesting observe that each basis element $\beta_j \in B$ can be written as a real $N$-th root of some element in $F$. In fact, $\beta_j$ is a radical over $F$ and generates a subfield of $E$ over $F$. Hence its degree $N_j$ over $F$ must be a divisor of $N$.

By Theorem 3.3, p. 24, in this case $\beta_j$ can be written as a real root $\sqrt[N_j]{\omega_j}$, $\omega_j \in F$. Since $N_j | N \sqrt[1/d]{\gamma}$ can be replaced by $\sqrt[N_j]{\omega_j}$ for an appropriate $\omega_j \in F$. This finally proves the theorem.

In view of the last fact mentioned in the proof, Theorem 6.1.1 can be interpreted as saying that only $dN$-th roots help in denesting $\sqrt[1/d]{\gamma}$ and that $\sqrt[1/d]{\gamma}$ can be denested by real radicals if and only if an element $\gamma_0 \in F$ exists such that for the real roots $\sqrt[1/d]{\gamma_0}$ of $\gamma_0 \in \sqrt[1/d]{\gamma} \in F$. As it turns out, the slightly more precise description in Theorem 6.1.1 yields a more efficient algorithm.

In case $E$ is generated by a single square root $\sqrt{\rho}$ the basis $B$ consists of $1$ and $\sqrt{\rho}$ only. So in this case we immediately get the first part of the theorem of Borodin et al. mentioned in the introduction to this section.

Next observe that, given the equation $\gamma_0 \beta_j \gamma = \eta^d$, $\eta \in E$, we can apply the distinct field embeddings $\sigma_i$, $i = 0, 1, \ldots, N - 1$, of $E$ over $F$ to this equation. This yields $\gamma_0 \sigma_i(\beta_j) \sigma_i(\gamma) = \sigma_i(\eta)^d$, for all $i$.

Hence any $d$-th root of $\sigma_i(\gamma)$ can be written as a product of an appropriate $d$-th root of $\sigma_i(\beta_j)^{-1}$, an appropriate $d$-th root of $\gamma_0^{-1}$, and $\sigma_i(\eta)$. In general, the $d$-roots will be complex.

Recall from Section 2 that $\sigma_i(\gamma)$ is a root of the minimal polynomial of $\gamma$. Hence the different $d$-th roots of $\sigma_i(\gamma)$, $i = 0, 1, \ldots, N - 1$, are the roots of the minimal polynomial of $\sqrt[1/d]{\gamma}$, provided $\sqrt[1/d]{\gamma}$ is of degree $d$. By Siegel’s
Theorem 3.9, p. 28, the radicals appearing in the expression for $\gamma$ and $\eta$ will be mapped by each $\sigma_i$ onto certain complex radicals. Hence the equation $\gamma_0 \beta_j \gamma = d$ leads to a denesting for all conjugates of $\sqrt[4]{\gamma}$ over $F$. In [HH] this has been called an exact denesting.

Before we proceed with linear combinations of nested radicals and complex nested radicals we show that no sum of real depth one radicals over $F$ denesting $\sqrt[4]{\gamma}$ can consist of fewer terms than the one described by Theorem 6.1.1. We will even show that the denesting is basically unique.

Observe that the radicals appearing in the denesting of $\sqrt[4]{\gamma}$ following from Theorem 6.1.1 are linearly independent. In fact, they all differ from elements in $B$ only by the factor $(\sqrt[4]{\gamma_0} \sqrt[4]{\beta_j})^{-1}$.

Suppose that $\sqrt[4]{\gamma}$ can be denested in two different ways by sums of real radicals over $F$,

$$\sqrt[4]{\gamma} = \sum_{i=1}^{n} \kappa_i \sqrt[4]{\eta_i}, \ \kappa_i \in F, \ \kappa_i \neq 0,$$

and

$$\sqrt[4]{\gamma} = \sum_{j=1}^{n'} \lambda_j \sqrt[4]{\mu_j}, \ \lambda_j \in F, \ \lambda_j \neq 0.$$

We may assume that the $\sqrt[4]{\eta_i}$’s and $\sqrt[4]{\mu_j}$’s separately are linearly independent. Hence (see Corollary 3.10, p. 28)

$$\frac{\sqrt[4]{\eta_i}}{\sqrt[4]{\eta_m}} \notin F, \ \frac{\sqrt[4]{\mu_i}}{\sqrt[4]{\mu_m}} \notin F$$

for all pairs of different indices.

On the other hand,

$$\sum_{i=1}^{n} \kappa_i \sqrt[4]{\eta_i} - \sum_{j=1}^{n'} \lambda_j \sqrt[4]{\mu_j} = 0.$$

Therefore the set of radicals \{$\sqrt[4]{\eta_1}, \ldots, \sqrt[4]{\eta_n}, \sqrt[4]{\mu_1}, \ldots, \sqrt[4]{\mu_{n'}}$\} must be linearly dependent. It follows from Corollary 3.10, p. 28, and the fact that the $\sqrt[4]{\eta_i}$’s, $\sqrt[4]{\mu_j}$’s separately are linearly independent that any $\sqrt[4]{\eta_i}$ differs from some $\sqrt[4]{\mu_j}$ only by a factor in $F$. Moreover, different $\sqrt[4]{\eta_i}$’s are multiples of different $\sqrt[4]{\mu_j}$’s. The same is true if we interchange the roles of the $\sqrt[4]{\eta_i}$’s and $\sqrt[4]{\mu_j}$’s. This shows that up to linear depend radicals denestings of $\sqrt[4]{\gamma}$ by sums of real depth one radicals are basically unique and will have at most $N$ terms.
Observe that the number of terms in the description of \( \gamma \) may be less than \( \log N \). In general, elements of \( E \) have a description of size \( \Omega(N) \) but although we treat \( \gamma \) as an element of \( E \) we do not assume that \( E \) is fixed and that \( \gamma \) is given as an element of \( E \), that is, as a linear combination of elements in the standard basis \( B \). Rather in the algorithms we describe below we first determine an appropriate representation of \( E \). So if we really want to compute the denesting of Theorem 6.1.1 we must expect the output size to be exponential in the input size.

If we could represent elements in \( E \) not only as linear combinations of basis elements but by arbitrary rational expressions in certain elements of \( E \) the output size might be much smaller than \( N \). But as all other methods in algorithmic algebra our approach relies on the basis representation.

Also in special cases the output size with respect to a basis representation may be much smaller than \( N \) but it looks almost impossible to predict the output size in advance or to get an output sensitive algorithm. Hence we should be satisfied with a denesting algorithm whose run time is polynomial in \( N \). At the end of Section 6 we describe an algorithm that achieves such a run time.

As the next theorem shows denesting radicals of the form \( \sqrt[\gamma]{\cdot} \) already suffices to denest linear combinations of nested radicals.

**Theorem 6.1.2** Suppose \( S = \sum_{i=1}^{k} \kappa_i \sqrt[\gamma_i]{\cdot} \) is a sum of real nested radicals such that \( \kappa_i, \gamma_i \in E_i, i = 1, \ldots, k \), and each \( E_i \) is a real radical extension of \( F \subseteq \mathbb{R} \). Furthermore assume that no nested radical \( \sqrt[\gamma_i]{\cdot} \) denests using real radicals over \( F \). Then

- If \( S \) denests using real radicals then \( S = 0 \).
- If no quotient \( \frac{\sqrt[\gamma_i]{\cdot}}{\sqrt[\gamma_j]{\cdot}} \) \( i \neq j \), denests using real radicals then \( S = 0 \) if and only if \( \kappa_i = 0 \) for all \( i \).

**Proof:** If a quotient \( \frac{\sqrt[\gamma_i]{\cdot}}{\sqrt[\gamma_j]{\cdot}} \) denests then \( \frac{\sqrt[\gamma_i]{\cdot}}{\sqrt[\gamma_j]{\cdot}} = \sum_{h=1}^{l} \lambda_h \sqrt[\rho_h]{\cdot} \) for some \( \lambda_h, \rho_h \in F \). Hence \( S \) can be written as \( S = \sum_{i \in I} \kappa'_i \sqrt[\rho_i]{\cdot} \), where each \( \kappa'_i \) is a real depth 1 expression over \( F, I \subseteq \{1, \ldots, k\} \), and no quotient \( \frac{\sqrt[\gamma_i]{\cdot}}{\sqrt[\gamma_j]{\cdot}}, i, j \in I \) denests.

Next suppose \( S \) denests to \( S = \sum_{j=1}^{k'} \lambda_j \sqrt[\rho_j]{\cdot}, \lambda_j, \rho_j \in F \). Consider the real radical extension \( E \) of \( F \) generated by the fields \( E_i \), the radicals \( \sqrt[\rho_j]{\cdot} \),
and the radicals appearing in at least one of the $\kappa_i$’s.

$$S' = S - \sum_{j=1}^{k'} \lambda_j \sqrt[p]{\rho_j} = \sum_{i \in I} \kappa_i \sqrt[\gamma_i]{\gamma_i} - \sum_{j=1}^{k'} \lambda_j \sqrt[p]{\rho_j} = 0.$$  

Consider the sum $S'$ as a linear combination of the radicals $1, \sqrt[\gamma_i]{\gamma_i}, i \in I,$ where the coefficient $\kappa'_0$ for 1 is defined as $\kappa'_0 = \sum_{j=1}^{k'} \lambda_j \sqrt[p]{\rho_j}.$ Each radical $\sqrt[\gamma_i]{\gamma_i}$ is a depth one radical over $E$ and the same is trivially true for the element 1. Furthermore, the coefficients $\kappa'_i$ in $S'$ are elements of $E.$

Apply Corollary 3.10, p. 28, to the field $E$ and the sum $S'.$ Together with the assumption that no nested radical $\sqrt[\gamma_i]{\gamma_i}$ and the fact that no quotient of these nested radicals denests using real radicals it implies that the set $\{ \sqrt[\gamma_i]{\gamma_i} | i \in I \} \cup \{ 1 \}$ is linearly independent over $E$ and hence $S' = 0$ if and only if $\kappa'_i = 0$ for all $i.$ In particular, $S = \kappa'_0 = 0,$ which proves the first part of the theorem.

If the nested radicals in $S$ have already the property that no quotient of two of them denests then the argument above shows that $S = 0$ implies $\kappa_i = 0,$ for all $i,$ which proves the second part of the theorem.

With respect to this theorem, given a sum $S = \sum_{i=1}^{k} \kappa_i \sqrt[\gamma_i]{\gamma_i}$ of nested radicals, by determining which nested radicals and which ratios of nested radicals denest $S$ can be transformed into

$$S = S' + \sum_{j=1}^{k'} \lambda_j \sqrt[p]{\rho_j},$$

where $S'$ is a sum of nested radicals satisfying the conditions of the previous theorem and $\sum_{j=1}^{k'} \lambda_j \sqrt[p]{\rho_j}$ is a depth 1 expression. Hence $S$ can denest if and only if $S' = 0,$ in which case the denesting is given by $\sum_{j=1}^{k'} \lambda_j \sqrt[p]{\rho_j}.$

As above it can easily be shown that a denesting for $S$ leads to a denesting for all roots of the minimal polynomial of $S.$

Notice that the degree of the minimal polynomial of the sum $S$ is in general exponential in $k.$ On the other hand, from the theorem above it is easily deduced that the number of terms in a denesting is polynomial in this parameter. Moreover, we will prove below that the whole description size of the denesting is polynomial in $k.$ Hence the run times of the algorithms in [La2] and [HH] that construct the minimal polynomial and even the splitting
field of $S$ will be exponential in the output size. Our denesting algorithm for these denesting will have a run time that is polynomial in $k$.

For complex radicals we get the following generalization of Theorem 6.1.1. So far it is computationally not very useful but it partially proves a conjecture of R. Zippel [Z].

**Theorem 6.1.3** Assume $F$ contains all roots of unity. Let $E$ be an arbitrary radical extension of $F$ with standard basis $B$. Assume $\gamma \in E$ and $d \in \mathbb{N}$ such that $\sqrt[d]{\gamma}$ denests over $F$, that is, $\sqrt[d]{\gamma} \in F(\sqrt[d]{\gamma_1}, \ldots, \sqrt[d]{\gamma_m})$, for some $\gamma_i \in F$. Then there exist $\gamma_0 \in F \setminus \{0\}$, $\beta_j \in B$ such that

$$\sqrt[d]{\gamma_0} \sqrt[d]{\gamma_j} = \eta \in E$$

for all $d$-th roots of $\gamma_0, \beta_j$.

The proof of this theorem is exactly the same as for Theorem 6.1.1 except that Lemma 3.6, p. 26, is replaced by Lemma 3.8, p. 27 and that we need not worry that all radicals involved are real. The fact that we may choose arbitrary roots of $\gamma_0$ and $\beta_j$ is an immediate consequence of the fact that $F$ contains all roots of unity.

In the form presented above Theorem 6.1.3 seems to be of theoretical interest only since it is computationally infeasible to work with a field containing all roots of unity. But at least it shows that even for arbitrary nested radicals those denestings that will be considered in the remaining sections of this thesis may be the right thing to look at. Also, the results of Horng, Huang and Landau can be interpreted as saying that instead of considering a field containing all roots of unity it suffices to work in a field that contains a certain root of unity whose degree is finite although very large. Moreover, there is the following restricted version of Theorem 6.1.3. It also generalizes a result of Borodin et al. for certain complex nested square roots (see [BFHT]).

**Theorem 6.1.4** Assume $F$ contains a primitive $d$-th root of unity. Let $E$ be a radical extension of $F$ generated by radicals $\sqrt[d]{\rho_i}$ such that $d_i$ divides $d$. Assume $\gamma \in E$, $d \in \mathbb{N}$, such that $\sqrt[d]{\gamma}$ denests over $F$ using radicals of the form $\sqrt[t]{\rho}$ with $t$ dividing $d$, that is, $\sqrt[d]{\gamma} \in F(\sqrt[d_1]{\gamma_1}, \ldots, \sqrt[d_m]{\gamma_m})$, for some $\gamma_i \in F$, $t_i \in \mathbb{N}$, $t_i$ dividing $d$ for all $i$. Then there exists an element $\gamma_0 \in F \setminus \{0\}$ such that

$$\sqrt[d]{\gamma_0} \sqrt[d]{\gamma} = \eta \in E$$

for all $d$-th roots of $\gamma_0$. 102
Representing $\eta$ as a linear combination of the elements of the standard basis of $E$ and dividing it by $\sqrt[d]{\gamma_0}$ leads to a denesting of $\sqrt[2]{\gamma}$ using only radicals of the form $\sqrt[2]{\rho}$ with $t$ dividing $d$ and $\rho \in F$.

**Proof:** $\sqrt[2]{\gamma}$ is a radical over $E$ contained in $E(\sqrt[2]{\gamma_1}, \ldots, \sqrt[2]{\gamma_m})$. Since $t_i$ divides $d$, $i = 1, 2, \ldots, m$, the field $F$ and hence $E$ contains primitive $d$-th and primitive $t_i$-th roots of unity for all $i$. Applying Lemma 3.8, p. 27, shows

$$\sqrt[2]{\gamma} = \eta \prod_{i=1}^m \sqrt[2]{\gamma_i^{f_i}}, \text{ or } \sqrt[2]{\gamma} \left( \prod_{i=1}^m \sqrt[2]{\gamma_i^{f_i}} \right)^{-1} = \eta$$

for some $\eta \in E$, $f_i \in \mathbb{N}$.

Since $t_i$ divides $d$, the product $\left( \prod_{i=1}^m \sqrt[2]{\gamma_i^{f_i}} \right)^{-1}$ can be written as $d$-th root of $\gamma_0$ for an appropriate $\gamma_0 \in F$ and an appropriate $d$-th root of $\gamma_0$.

But $E$ contains all $d$-th roots of unity. Therefore multiplying $\sqrt[2]{\gamma}$ with any $d$-th roots of $\gamma_0$ leads to an element of $E$.

Since the extension $E$ has a standard basis consisting of radicals of the form $\sqrt[d]{\rho}$, $t$ dividing $d$, the equation $\sqrt[d]{\gamma_0} \sqrt[2]{\gamma} = \eta \in E$ really leads to a denesting of $\sqrt[2]{\gamma}$ using only radicals of the form $\sqrt[d]{\rho}$, $\rho \in F$, $t \in \mathbb{N}$, such that $t$ divides $d$.

We may have formulated the theorem using only radicals of the form $\sqrt[d]{\rho}$, $\rho \in F$. However, the present form contains more information and is more convenient for the analysis of the denesting algorithms in Section 7.

The theorem may also be interesting from a practical point of view in the following sense.

Given a nested radical like $\sqrt[2]{\gamma}$ containing complex radicals $\sqrt[2]{\rho_i}$, assume you want to denest it and you want to allow a denesting by radicals over some larger field, which is still easy to describe and allows efficient arithmetic, and you want to allow the degrees of the radicals to be larger than the original ones.

This can be achieved in the following way. Choose some integer $D$ such that $d$ and $d_i$ divide $D$. $\sqrt[2]{\gamma}$ can be written as $\sqrt[d]{\gamma}$ for some $e$. Replace the field $F$ by the smallest field $F'$ containing $F$ and a primitive $D$-th root of unity. Then Theorem 6.1.4 describes and the algorithm we are going to describe in the remainder of Section 6 and in Section 7 can determine a denesting over $F'$ using radicals of the form $\sqrt[d]{\rho}$, $\rho \in F'$.
This generalization is easily incorporated in the results and techniques described below. It does not require any new arguments or ideas. So we will restrict ourselves in the sequel to denestings as in Theorem 6.1.4.

Recall that in the real case we showed that the denesting of Theorem 6.1.1 was basically unique and since the elements in the standard basis of a radical extension are of course linearly independent any denesting of $\sqrt[d]{\gamma}$ by a sum of real radicals has at least as many terms as the one obtained by Theorem 6.1.1. The same is true in the present situation if the expression “real radicals” is always replaced by “radicals of the form $\sqrt[t]{\rho}$ with $t$ dividing $d$”. In particular, any denesting of this form has at least as many terms as the one obtained by the conclusion in Theorem 6.1.4.

Let us finally mention the generalization of Theorem 6.1.2 to complex radicals.

**Theorem 6.1.5** Assume $F$ contains a primitive $d$-th root of unity. Suppose $S = \sum_{i=1}^{k} \kappa_i \sqrt[d]{\gamma_i}$, is a sum of nested radicals such that $\kappa_i, \gamma_i \in E_i$, $i = 1, \ldots, k$, and each $E_i$ is a radical extension of $F$ generated by radicals of the form $\sqrt[t]{\rho}$, $\rho \in F$, $t \in \mathbb{N}$, such that $t$ divides $d$. Furthermore assume that no nested radical $\sqrt[d]{\gamma_i}$ denests using only radicals of the same form as the ones defining the fields $E_i$. Then

- If $S$ denests using radicals of the form $\sqrt[t]{\rho}$, $\rho \in F$, $t|d$, then $S = 0$
- If no quotient $\frac{\sqrt[d]{\gamma_i}}{\sqrt[d]{\gamma_j}}$, $i \neq j$, denests using radicals of the form $\sqrt[t]{\rho}$, $\rho \in F$, $t|d$, then $S = 0$ if and only if $\kappa_i = 0$ for all $i$.

The proof is as the one for Theorem 6.1.2 replacing the real variant of Corollary 3.10, p. 28, by the complex one.

This finishes the description of our basic denesting theorems. In the following section we give a description of all elements $\gamma_0 \in F$ with the property $\sqrt[d]{\gamma_0} \sqrt[d]{\gamma} \in E$ for a nested radical $\sqrt[d]{\gamma}$, $\gamma \in E$. This is already an exhaustive description of denestings as in Theorem 6.1.4. To describe the denestings of Theorem 6.1.1 in full detail we apply the description to all elements $\beta_j \gamma$, $\gamma \in E$, $\beta_j \in B$.

This description, however, leads to an efficient algorithm only for nested radicals $\sqrt[d]{\gamma}$, where $\gamma$ is an element of a simple radical extension. But as it turns out, any denesting problem can be reduced to certain denesting problems for simple radical extensions.
6.2 Denesting Sets and Reduction to Simple Radical Extensions

With respect to the results of Theorem 6.1.1 and Theorem 6.1.4 for the rest of Section 6 we use the following convention.

**Convention** For the rest of Section 6 any symbol of the form $m \sqrt[\rho]{\gamma}$, $m \in \mathbb{N}$, $\rho \in \mathbb{C}$, denotes the complex number $|\rho|^{1/m} \left( \cos \frac{1}{m} \phi_{\rho} + i \sin \frac{1}{m} \phi_{\rho} \right)$, where $|\rho|^{1/m}$ is the positive real $m$-th root of $|\rho|$ and $\phi_{\rho} \in (-\pi, \pi]$ is the angle of $\rho$ when written in polar coordinates.

In fact, in Theorem 6.1.4 since $F$ contains a primitive $d$-th root of unity and $d_i | d$ for all $i$ the extension $F(\sqrt[d_1]{\rho_1}, \ldots, \sqrt[d_k]{\rho_k})$ is the same no matter which $d_i$-th roots are meant. Furthermore, whether the nested radical can be denested depends only on $d$ and $\gamma$ but not on the particular choice of the $d$-th root. Finally, we may choose the roots of the element $\gamma_0$ in Theorem 6.1.4 arbitrarily.

Likewise, in Theorem 6.1.1 the radical extension $F(\sqrt[d]{\rho_1}, \ldots, \sqrt[d]{\rho_k})$ is independent of the specific real roots and the question whether the nested radical can be denested depends only on $d, \gamma$, and the fact that real $d$-th roots are considered. Moreover, if elements $\gamma_0, \beta_j$ as stated in the theorem exist we may take any real $d$-th root of $\gamma_0, \beta_j$.

Observe that if a real number $\rho$ has a real $m$-th root then $|\rho|^{1/m} \left( \cos \frac{1}{m} \phi_{\rho} + i \sin \frac{1}{m} \phi_{\rho} \right)$ uniquely describes such a real root. Finally, if $\rho > 0$ then $\sqrt[m]{\rho} > 0$, too.

Also to simplify the statement of our results we refer to denesting problems as in Theorem 6.1.1 and Theorem 6.1.2 as the real case. To denesting problems as in Theorem 6.1.4 and Theorem 6.1.5 we refer as the complex case.

**Definition 6.2.1** Assume $F \subset \mathbb{C}$ and let $E$ be an algebraic extension of $F$, $\gamma \in E$, $d \in \mathbb{N}$. We say that $\gamma_0 \in F \setminus \{0\}$ denests $\sqrt[\gamma]{\gamma}$ over $F$ or that it leads to a denesting of $\sqrt[\gamma]{\gamma}$ over $F$ if $\gamma_0 \gamma = \eta^d$ for some $\eta \in E$. The element $\gamma_0$ is called a denesting element for $\sqrt[\gamma]{\gamma}$ over $F$.

A set $S \subset F$ of elements denesting $\sqrt[\gamma]{\gamma}$ such that any element denesting $\sqrt[\gamma]{\gamma}$ differs from some element in this set by a $d$-th power of an element in $F$ is called a denesting set for $\sqrt[\gamma]{\gamma}$ over $F$. If no two elements in a denestings set differ from one another by a $d$-th power of an element in $F$ then the denesting set is called irreducible.
We gave this definition for arbitrary fields $F$ and $E$ because our characterization for elements denesting a nested radical we give below applies in such a general setting and therefore generalizes a characterization of Zippel [Z] and Landau [La3].

In the general setting the definition says that $\gamma_0$ denests $\sqrt[d]{\gamma}$ if $\zeta \sqrt[d]{\gamma_0} \sqrt[d]{\gamma} \in E$ for some $d$-th root of unity $\zeta$. By the remarks above for field extensions as in Theorem 6.1.1 or Theorem 6.1.4 it simply states that $\gamma_0$ denests $\sqrt[d]{\gamma}$ if and only if $\sqrt[d]{\gamma_0} \sqrt[d]{\gamma} \in E$ (recall the convention from the beginning of this section). Notice that in the real case if $d$ is even then $\gamma_0$ can only denest $\sqrt[d]{\gamma}$ only if $\gamma_0$ is positive.

Before we characterize denesting elements we show that any denesting problem as in Theorem 6.1.1 or in Theorem 6.1.4 can be reduced via denesting sets to certain denesting problems over simple radical extensions.

We begin with showing that in these cases finite denesting sets exist.

**Lemma 6.2.2** Let the fields $F, E, d,$ and $\gamma$ be either as in Theorem 6.1.1 or as in Theorem 6.1.4. By $B$ denote the standard basis of the extension $E$. If $\gamma_0 \in F$ denests $\sqrt[d]{\gamma}$ then

\[
\{ \beta_j \gamma_0 | \beta_j \in B \text{ such that } \beta_j^d \in F \}
\]

and

\[
\{ \beta_j^{-d} \gamma_0 | \beta_j \in B \text{ such that } \beta_j^d \in F \}
\]

are irreducible denesting sets for $\sqrt[d]{\gamma}$ over $F$.

The second set has been included since it turns out to be more convenient from a computational point of view.

**Proof:** It is clear that both sets contain only denesting elements for $\sqrt[d]{\gamma}$. It remains to show that they are irreducible denesting sets. We will do so in detail only for the first set.

First let us show that any two elements $\gamma_0, \gamma_0'$ that denest $\sqrt[d]{\gamma}$ differ from one another by a factor of the form $\rho \beta_j^d$, with $\rho \in F, \beta_j \in F$ such that $\beta_j^d \in B$. This will show that the first set is a denesting set.

By the remarks above

$\sqrt[d]{\gamma_0} \sqrt[d]{\gamma} \in E,$

and, likewise,

$\sqrt[d]{\gamma_0} \sqrt[d]{\gamma} \in E.$
Hence
\[
\frac{\sqrt[d]{\gamma_0}}{\sqrt[d]{\gamma}} = \frac{\sqrt[d]{\gamma_0}}{\sqrt[d]{\gamma}} \in E.
\]
Therefore the ratio is a radical of $E$ over $F$. In particular, if $E$ is real the ratio must be a real number. By Lemma 3.6, p. 26 in the real case, or Lemma 3.8, p. 27 in the complex case,
\[
\frac{\sqrt[d]{\gamma_0}}{\sqrt[d]{\gamma}} = \rho \beta_j,
\]
for some $\rho \in F$, $\beta_j \in B$. The claim follows.

It remains to show that the set is irreducible. If the elements in the set did not form an irreducible denesting set then
\[
\frac{\beta_i^{d_i} \gamma_0}{\beta_j^{d_j} \gamma_0} = \rho^d
\]
for some $\rho \in F$ and different basis elements $\beta_i, \beta_j$. Hence $\beta_j$ and $\beta_i$ would differ from one another only by a factor in $F$. This is impossible since $\beta_i$ and $\beta_j$ are linearly independent.

The claim for the second set of elements in $F$ can be shown in exactly the same way by observing that if $E$ is described as the extension generated by $\sqrt[d]{\rho_i^{-1}}$ then with respect to these generators the elements $\beta_j^{-1}$ form the standard basis.

In the complex case $\beta_j^d \in F$ for all elements $\beta_j$ of the standard basis $B$ hence the previous lemma implies

**Corollary 6.2.3** In the real case a denesting set has size at most $N$. In the complex case a denesting set has size exactly $N$.

Assume next that $L \supset F$ is a proper subfield of $E$, where $F$ and $E$ are as in the previous lemma. We may consider $\gamma$ also as an element of the algebraic extension $E$ of the field $L$. If an element in $F$ denesting $\sqrt[d]{\gamma}$ over $F$ exists then it is also a denesting element in $L$ for $\sqrt[d]{\gamma}$ over $L$. Moreover, let $D$ be a set containing the inverses of the elements of a denesting set for $\sqrt[d]{\gamma}$ over $L$. For each element $\gamma' \in D$ we consider a denesting set of $\sqrt[d]{\gamma'}$ over $F$. We claim that the union of these sets contains a denesting set for $\sqrt[d]{\gamma}$ over $F$.

107
To prove the claim recall that any element $\gamma_0 \in F \subset L$ that denests $\sqrt[d]{\gamma}$ over $F$ also denests $\sqrt[d]{\gamma}$ over $L$. Hence it differs from the inverse of some element $\gamma' \in D$ by a $d$-th power of an element in $L$, i.e.,

$$\gamma_0 = \eta^d \frac{1}{\gamma'}, \eta \in L.$$ 

Hence

$$\gamma_0 \gamma' = \eta^d,$$

i.e., $\gamma_0$ is a denestiong element for $\sqrt[d]{\gamma}$. By definition of a denesting set the claim follows.

Denote by $F^{(i)}$ the subfield of $E$ generated by the first $i$ radicals $\sqrt[d]{\rho_1}, \ldots, \sqrt[d]{\rho_i}$. Define sets $D^{(k-i)}$, $i = 0, 1, \ldots, k$, as follows:

$$D^{(k)} = \{\gamma\}.$$ 

Assume $D^{(k-i)}$ has already been defined. For each element $\gamma^{(k-i)} \in D^{(k-i)}$ let $D_{\gamma^{(k-i)}}$ be a set containing the inverses of a denesting set for $\sqrt[d]{\gamma^{(k-i)}}$ over $F^{(k-i-1)}$. Then

$$D^{(k-i-1)} = \bigcup_{\gamma^{(k-i)} \in D^{(k-i)}} D_{\gamma^{(k-i)}}.$$ 

By induction on $k - i$ and by repeatedly using the argument from above we immediately get

Lemma 6.2.4 If the sets $D^{(k-i)}$ are defined as above then the inverses of the elements in $D^{(0)}$ form a superset of a denesting set for $\sqrt[d]{\gamma}$.

Our basic denesting algorithm will compute a set $D^{(0)}$ as described above and then will check for all inverses of elements in $D^{(0)}$ whether they denest $\sqrt[d]{\gamma}$.

But as long as we are not able to characterize and compute denesting elements the results obtained above are of no use. Therefore in the next subsection we give a constructive description of the elements denesting $\sqrt[d]{\gamma}$. It applies not only to radical extensions but to arbitrary extensions.

6.3 Characterizing Denesting Elements

In this subsection let $F \subset C$ be an arbitrary field and $E$ an algebraic extension of $F$. 

108
We need one more definition. Let $\gamma \in E$ and $d \in \mathbb{N}$. If $\gamma_0$ denests $\sqrt[\gamma]{\gamma}$ then $\gamma_0 \gamma = \eta^d$, $\eta \in E$. Recall that this implies $\gamma_0 \sigma_i(\gamma) = \sigma_i(\eta)^d$ for all embeddings $\sigma_i$ of $E$ over $F$. These equations are equivalent to the existence of $d$-th roots of unity $\zeta^{(i)}$ such that

$$\zeta^{(i)} \sqrt[\gamma_0]{d} \sqrt[\sigma_i(\gamma)]{\gamma} = \sigma_i(\eta), \text{ for all } i.$$ 

**Definition 6.3.1** Let $F$ be an arbitrary field, $E$ an extension of $F$, $[E : F] = N$, $\gamma \in E, d \in \mathbb{N}$. Denote the field embeddings of $E$ over $F$ by $\sigma_i$, $i = 0, 1, \ldots, N - 1$, with $\sigma_0$ being the identity. A sequence $(\zeta^{(0)}, \zeta^{(1)}, \ldots, \zeta^{(N-1)})$ of $d$-th roots of unity is called an **admissible sequence** for $\sqrt[\gamma]{\gamma}$ over $F$ if elements $\gamma_0 \in F$ and $\eta \in E$ exist such that $\zeta^{(i)} \sqrt[\gamma_0]{d} \sqrt[\sigma_i(\gamma)]{\gamma} = \sigma_i(\eta)$ for all $i = 0, 1, \ldots, N - 1$. For fixed $\gamma_0 \in F$ such a sequence is called an admissible sequence **corresponding to** $\gamma_0$.

An admissible sequence is called **normalized** if $\zeta^{(0)} = 1$.

For $\gamma_0 \in F$ there may be many different admissible sequences corresponding to it. For example, assume that $F$ contains a primitive $d$-th root of unity. If at least one admissible sequence corresponding to $\gamma_0$ exists then there are exactly $d$ different admissible sequences corresponding to $\gamma_0$. On the other hand, since we fixed the meaning of $\sqrt[\gamma_0]{d} \sqrt[\gamma]{\gamma}$, for each $\gamma_0$ there is at most one normalized sequence corresponding to it.

Based on this definition we give our characterization of elements denesting $\gamma$. It generalizes results of R. Zippel [Z] and S. Landau [La3] who characterized elements denesting a $d$-th root of unity if $E$ is a Galois extension of $F$. Lemma 6.3.2 avoids this assumption and is correct for arbitrary extensions of an arbitrary field $F$.

**Lemma 6.3.2** Let $F \subset C$ be an arbitrary field and $E$ an algebraic extension of $F$ with basis $\{\beta_0, \beta_1, \ldots, \beta_{N-1}\}$. Denote the distinct field embeddings by $\sigma_i$, $i = 0, \ldots, N - 1$, with $\sigma_0$ being the identity. Assume $\gamma \in E$, $d \in \mathbb{N}$. If a denesting element for $\sqrt[\gamma]{\gamma}$ exists then an admissible sequence $(\zeta^{(0)}, \zeta^{(1)}, \ldots, \zeta^{(N-1)})$ for $\sqrt[\gamma]{\gamma}$ and a basis element $\beta_j$ exist such that

$$\left(\sum_{i=0}^{N-1} \zeta^{(i)} \sigma_i(\beta_j) \sqrt[\sigma_i(\gamma)]{\gamma}\right)^d \in F\{0\}$$

and the inverse of this element also denests $\sqrt[\gamma]{\gamma}$ over $F$. 

109
Moreover, any element $\gamma_0 \in F$ that denests $\sqrt[d]{\gamma}$ can be written as

$$\rho^d \left( \sum_{i=0}^{N-1} \zeta^{(i)} \sigma_i(\beta_j) \sqrt[d]{\sigma_i(\gamma)} \right)^{-d}$$

for an admissible sequence $(\zeta^{(0)}, \zeta^{(1)}, \ldots, \zeta^{(N-1)})$ for $\sqrt[d]{\gamma}$, a basis element $\beta_j$, and some $\rho \in F$.

**Proof:** Assume $\gamma_0 \in F$ denests $\sqrt[d]{\gamma}$. Hence an $\eta \in E$ exists such that $\gamma_0 \sigma_i(\gamma) = \sigma_i(\eta)^d$ for all embeddings $\sigma_i$. Equivalently, $d$-th roots of unity $\zeta^{(i)}$ exist such that $\zeta^{(i)} \sqrt[d]{\gamma_0} \sqrt[d]{\gamma} = \sigma_i(\eta)$.

These roots of unity form an admissible sequence.

Furthermore applying Lemma 3.5, p. 25, to $\eta \in E$ yields

$$\text{tr}(\beta_j \eta) = \sum_{i=0}^{N-1} \sigma_i(\beta_j) \sigma_i(\eta) = \sqrt[d]{\gamma_0} \sum_{i=0}^{N-1} \zeta^{(i)} \sigma_i(\beta_j) \sqrt[d]{\sigma_i(\gamma)} = \rho \in F \setminus \{0\}$$

for some element $\beta_j$ of the basis. Therefore

$$\gamma_0 = \left( \sum_{i=0}^{N-1} \zeta^{(i)} \sigma_i(\beta_j) \sqrt[d]{\sigma_i(\gamma)} \right)^d.$$

$\rho \in F$ hence $\left( \sum_{i=0}^{N-1} \zeta^{(i)} \sigma_i(\beta_j) \sqrt[d]{\sigma_i(\gamma)} \right)^d$ is also an element of $F$.

Moreover, as a $d$-th power of an element in $F \subset E$ the element $\rho^d$ will not help in denesting $\sqrt[d]{\gamma}$, so $\left( \sum_{i=0}^{N-1} \zeta^{(i)} \sigma_i(\beta_j) \sqrt[d]{\sigma_i(\gamma)} \right)^{-d}$ will also denest $\sqrt[d]{\gamma}$. This proves the existence of an admissible sequence and basis element as stated.

The claim that any number $\gamma_0 \in F$ denesting $\sqrt[d]{\gamma}$ can be written in the stated form follows from the fact that the process we just described can be applied to all these elements of the field $F$.

The existence proof given above generalizes Lemma 6.2.2 in the sense that it shows the existence of finite denesting sets for expression $\sqrt[d]{\gamma}$ even if $F$ and $E$ are arbitrary fields.

**Remark 6.3.3** If $E$ and $\sqrt[d]{\gamma}$ are as in Theorem 6.1.1 or as in Theorem 6.1.4 then Lemma 6.3.2 remains correct if “admissible sequence” is replaced
by “normalized admissible sequence”. In fact, in the complex case, for the identity \( \sigma_0 \), that is for \( E \) and \( \gamma \) itself it does not matter which \( d \)-th root of \( \sqrt[d]{\gamma_0} \) we are taking since \( F \) contains all \( d \)-th roots of unity, and hence we can take \( \sqrt[d]{\gamma_0} \) itself. In the real case, for \( i = 0 \) in \( \zeta(0) \sqrt[2]{\gamma_0} \gamma = \eta \in E \) the root of unity \( \zeta(0) \) must be real since \( \eta, \sqrt[2]{\gamma_0}, \) and \( \sqrt[2]{\gamma} \) are real. But then it may very well be 1.

Let us briefly show that Lemma 6.3.2 generalizes the second part of the theorem of Borodin et al. (Theorem 6.2).

Suppose \( E = F(\sqrt[p]{\rho}) \) and \( \gamma = \alpha + \beta\sqrt[p]{\rho}, \alpha, \beta \neq 0 \). Moreover assume an element \( \gamma_0 \in F \) exists such that \( \sqrt[p]{\gamma_0} \sqrt[p]{\alpha + \beta\sqrt[p]{\rho}} \in F \).

\( E \) has only two field embeddings, the identity and the mapping \( \sigma \) satisfying \( \sigma(\sqrt[p]{\rho}) = -\sqrt[p]{\rho} \). It is easily seen that for any \( \gamma_0 \) denesting \( \sqrt[p]{\gamma} \) \[ \text{tr} \left( \sqrt[p]{\gamma_0}(\alpha + \beta\sqrt[p]{\rho}) \right) \neq 0. \]

By the previous lemma

\[ \left( \sqrt[p]{\alpha + \beta\sqrt[p]{\rho}} + \epsilon \sqrt[p]{\alpha - \beta\sqrt[p]{\rho}} \right)^2 \in F, \]

for \( \epsilon = 1 \) or \( \epsilon = -1 \).

\[ \left( \sqrt[p]{\alpha + \beta\sqrt[p]{\rho}} + \epsilon \sqrt[p]{\alpha - \beta\sqrt[p]{\rho}} \right)^2 = 2\alpha + 2\epsilon \sqrt[p]{\alpha^2 - \rho \beta^2}. \]

Hence \( \sqrt[p]{\alpha^2 - \rho \beta^2} \in F \). This shows that for both choices of \( \epsilon \) the square above is in \( F \).

Moreover,

\[ \frac{1}{2\alpha + 2\sqrt[p]{\alpha^2 - \rho \beta^2}} \]

denests \( \sqrt[p]{\gamma} \), so the same must be true for \( 2\alpha + 2\sqrt[p]{\alpha^2 - \rho \beta^2} \) itself, which is the first part of the condition in the theorem of Borodin et al.. The second part can be deduced similarly.

Unfortunately, for \( d = 2 \) and extensions generated by more than one square root or for \( d > 2 \) Lemma 6.3.2 does not lead to such a simple condition. We have to work much harder to get efficient algorithms. And even then it is not possible to apply these algorithms recursively to nested radicals of depth larger than two as is the case for nested radicals containing exactly one square root on each level (for the details of this recursive algorithm see [BFHT]).

On the other hand observe that even if we cannot determine the admissible sequences for \( \sqrt[p]{\gamma} \) Lemma 6.3.2 almost immediately leads to an
algorithm that computes a denesting element for $\sqrt[d]{\gamma}$ if any such element exists. For all basis elements $\beta_j$ and all $dN$-tuples of $d$-th roots of unity we can use the algorithm leading to Theorem 5.2.6, p. 61, to check whether $(\sum_{i=0}^{N-1} \zeta(i)^d \sigma_i(\beta_j) \sqrt[d]{\sigma_i(\gamma)})^d$ equals an element $\gamma_0$ in $F$. Then we check for all elements $\gamma_0$ of $F$ found in this way whether $\zeta^{(0)} \sqrt[d]{\gamma_0^{-1}} \gamma \in E$ again using the method of Theorem 5.2.6, p. 61.

For the general case we do not know any better way to compute a denesting element than this brute force method. In case $E$ is a simple radical extension, however, we can do better. In the next subsection we give a characterization of admissible sequences that allows us to determine a superset of the set of normalized admissible sequences of size at most $d^3$. Then the process sketched above needs to be applied only to these sequences.

6.4 Characterizing Admissible Sequences

Throughout this subsection let $F$ be a field and $E = F(\sqrt[d]{\rho})$ a simple radical extension of $F$ such that $\sqrt[d]{\rho}$ is of degree $m$ over $F$. As usual $\gamma$ is an element of $E$.

Before we can prove the main result on admissible sequences we need two auxiliary lemmata.

Lemma 6.4.1 If $(\zeta^{(0)}, \zeta^{(1)}, \zeta^{(2)}, \ldots, \zeta^{(m-1)})$ is an admissible sequence for $\sqrt[d]{\gamma}$ then for any two indices $i, j$

$$\zeta^{(i)} d^{-1} \sqrt[d]{\sigma_i(\gamma)} \zeta^{(j)} d^{-1} \sqrt[d]{\sigma_j(\gamma)} \in F(\sqrt[d]{\rho}, \zeta_m).$$

Here $\zeta_m$ denotes a primitive $m$-th root of unity.

Proof: Since $(\zeta^{(0)}, \zeta^{(1)}, \zeta^{(2)}, \ldots, \zeta^{(m-1)})$ is an admissible sequence there exist elements $\gamma_0 \in F$, $\eta \in F(\sqrt[d]{\rho})$ such that $\zeta^{(i)} \sqrt[d]{\gamma_0} \sqrt[d]{\sigma_i(\gamma)} = \sigma_i(\eta)$ for all $i = 0, 1, \ldots, m - 1$. Hence

$$\zeta^{(i)} d^{-1} \sqrt[d]{\gamma_0} \zeta^{(j)} d^{-1} \sqrt[d]{\sigma_i(\gamma)} \zeta^{(j)} d \sqrt[d]{\sigma_j(\gamma)} =$$

$$= \gamma_0 \zeta^{(i)} d^{-1} \sqrt[d]{\sigma_i(\gamma)} d^{-1} \zeta^{(j)} d \sqrt[d]{\sigma_j(\gamma)} \in F(\sigma_i(\eta), \sigma_j(\eta)).$$

Since $\gamma_0 \in F$ we get $\zeta^{(i)} d^{-1} \sqrt[d]{\sigma_i(\gamma)} d^{-1} \zeta^{(j)} d \sqrt[d]{\sigma_j(\gamma)} \in F(\sigma_i(\eta), \sigma_j(\eta))$, too.

Because the conjugate fields of $F(\sqrt[d]{\rho})$ are of the form $F(\zeta_m \sqrt[d]{\rho})$

$$F(\sigma_i(\eta), \sigma_j(\eta)) \subseteq F(\sqrt[d]{\rho}, \zeta_m)$$

112
and the lemma follows.

As before we are basically interested in two different cases. First, \( \zeta_m \in F \), equivalently, \( F(\zeta_m) = F \), where \( \zeta_m \) is a primitive \( m \)-th root of unity, and, second, \( F \subseteq \mathbb{R} \). In the latter case we need to determine the degree of \( \sqrt[m]{\rho} \) over \( F(\zeta_m) \). We show that this degree is either \( m \) or \( \frac{m^2}{2} \).

**Lemma 6.4.2** Let \( F \) be a real field. Suppose \( f(X) = X^m - \rho, \rho \in F, \rho > 0 \), is irreducible in \( F[X] \). Furthermore let \( \zeta_m \) be a primitive \( m \)-th root of unity. Then \( f(X) \) is either irreducible in \( F(\zeta_m)[X] \), or \( \sqrt[m]{\rho} \in F(\zeta_m) \) and \( f(X) \) factors as \( (X^{\frac{m}{2}} - \sqrt[m]{\rho})(X^{\frac{m}{2}} + \sqrt[m]{\rho}) \) over \( F(\zeta_m)[X] \). In the latter case \( m \) must be even.

**Proof:** Consider the degree of \( \sqrt[m]{\rho} \) over \( F(\zeta_m) \). By Theorem 3.3, p. 24, this is the smallest positive integer \( e \) such that \( m \sqrt[m]{\rho}^e \in F(\zeta_m) \). Moreover, \( e \) divides \( m \).

By Lemma 3.12, p. 32, if \( \sqrt[m]{\rho} \in F(\zeta_m) \) then it must be a square root of an element in \( F \). Therefore \( \sqrt[m]{\rho}^{2e} \in F \). Since we assume that the degree of \( \sqrt[m]{\rho} \) over \( F \) is \( m \) it follows \( m|2e \). But \( e|m \) hence \( \frac{m}{2} = 1 \) or \( \frac{m}{2} = 2 \), which proves the lemma.

As follows from the lemma whenever \( m \) is odd then the degree of \( \sqrt[m]{\rho} \) over \( F \) and the degree of \( \sqrt[m]{\rho} \) over \( F(\zeta_m) \) are the same. The following example shows that the second situation described by the lemma can also occur. Consider the positive 10-th root of 5. Its degree over \( \mathbb{Q} \) is 10. On the other hand, \( \mathbb{Q}(\zeta_{10}) \) contains the square roots of 5. Hence the degree of \( \sqrt[10]{5} \) over the 10-th cyclotomic field is 5. The fact that \( \mathbb{Q}(\zeta_{10}) \) contains the square roots of 5 follows from a general result in algebraic number theory but it can also be seen directly as follows.

The fifth cyclotomic field \( \mathbb{Q}(\zeta_5) \) is a subfield of \( \mathbb{Q}(\zeta_{10}) \). Since \( \zeta_5^2 + \zeta_5^3 + \zeta_5 + 1 = 0 \), the element \( \zeta_5 + \zeta_5^{-1} = \zeta_5 + \zeta_5^4 \) satisfies \( (\zeta_5 + \zeta_5^{-1})^2 + (\zeta_5 + \zeta_5^{-1}) - 1 = 0 \). Hence it is a root of \( X^2 + X - 1 \) and can be written as \( \frac{1}{2}(1 + \sqrt{5}) \) for the positive or negative square root of 5. For this square root, \( \sqrt{5} = 2(\zeta_5 + \zeta_5^{-1}) - 1 \in \mathbb{Q}(\zeta_5) \subset \mathbb{Q}(\zeta_{10}) \).

**Lemma 6.4.3** Let \( F \subseteq \mathbb{R}, \rho \in F, \rho > 0 \), \( \sqrt[m]{\rho} \in \mathbb{R}, \gamma \in F(\sqrt[m]{\rho}), d \in \mathbb{N} \), such that the degree of \( \sqrt[m]{\rho} \) over \( F \) is \( m \). Any two normalized admissible
sequences for $\sqrt[m]{\gamma}$ over $F$ that have a common prefix $(1, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)})$ are equal.

If $F \subset \mathbb{C}$ contains a primitive $m$-th root of unity and $\rho \in F$, $\gamma \in F(\sqrt[m]{\rho})$, $d \in \mathbb{N}$, then $\zeta^{(1)}$ alone determines a normalized admissible sequence for $\sqrt[m]{\gamma}$ over $F$.

**Proof:** We will prove the lemma first for a real field $F$.

Let $(1, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)})$ be a common prefix of two admissible sequences $(1, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)}, \ldots, \zeta^{(m-1)})$, $(1, \zeta^{(1)}', \zeta^{(2)}', \zeta^{(3)}', \ldots, \zeta^{(m-1)}')$, hence $\zeta^{(i)} = \zeta^{(i)}'$, $i = 1, 2, 3$.

First we show by induction on $l = 0, 1, \ldots, \lfloor \frac{m}{2} \rfloor$, that $\zeta^{(2l)} = \zeta^{(2l)}$.

By assumption, $\zeta^{(0)} = \zeta^{(0)}' = 1$, and $\zeta^{(2)} = \zeta^{(2)}'$. Now assume that $\zeta^{(2l)} = \zeta^{(2l)}'$ has already been shown for all $l = 0, 1, \ldots, h - 1$. By definition of admissible sequences elements $\rho_1, \rho_2 \in F$ and $\eta_1, \eta_2 \in F(\sqrt[m]{\rho})$ exist such that

$$\zeta^{(2(h-j))} \sqrt[m]{\rho_1} \sqrt[m]{\sigma_{2(h-j)}(\gamma)} = \sigma_{2(h-j)}(\eta_1),$$

$$\zeta^{(2(h-j))} \sqrt[m]{\rho_2} \sqrt[m]{\sigma_{2(h-j)}(\gamma)} = \sigma_{2(h-j)}(\eta_2), \text{ for } j = 1, 2,$$

and

$$\zeta^{(2k)} \sqrt[m]{\rho_1} \sqrt[m]{\sigma_{2k}(\gamma)} = \sigma_{2k}(\eta_1),$$

$$\zeta^{(2k)} \sqrt[m]{\rho_2} \sqrt[m]{\sigma_{2k}(\gamma)} = \sigma_{2k}(\eta_2).$$

W.l.o.g. we may assume that the isomorphism $\sigma_i$ of $F(\sqrt[m]{\rho})$ over $F$ is given by $\sigma_i(\sqrt[m]{\rho}) = \zeta_i \sqrt[m]{\rho}$, $i = 0, 1, \ldots, m - 1$, for some fixed primitive $m$-th root of unity $\zeta_m$.

By Lemma 6.4.1 the first equations imply

$$\frac{1}{\rho_1} \sigma_{2(h-1)}(\eta_1) \left( \sigma_{2(h-2)}(\eta_1) \right)^{d-1} =$$

$$= \zeta^{(2(h-1))} \zeta^{(2(h-2))} d^{-1} \sqrt[m]{\sigma_{2(h-1)}(\gamma)} \left( \sqrt[m]{\sigma_{2(h-2)}(\gamma)} \right)^{d-1} =$$

$$= \frac{1}{\rho_2} \sigma_{2(h-1)}(\eta_2) \left( \sigma_{2(h-2)}(\eta_2) \right)^{d-1} = \eta \in F(\zeta_m, \sqrt[m]{\rho}).$$

By the previous lemma a field isomorphism $\tau$ of $F(\zeta_m, \sqrt[m]{\rho})$ over $F(\zeta_m)$ exists that maps $\sqrt[m]{\rho}$ onto $\zeta_m^2 \sqrt[m]{\rho}$.

\[14\text{Recall that the degree of } \sqrt[m]{\rho} \text{ over } F \text{ is } m.\]
We want to apply \( \tau \) to the equations above. \( \sigma_i(\eta_1) \) and \( \sigma_i(\eta_2) \) are elements of \( F(\zeta_m, \sqrt{\rho}) \) for all embeddings \( \sigma_i \) and we only have to determine the images of \( \sigma_i(\eta_1), \sigma_i(\eta_2) \) under \( \tau \). If \( \eta_1 = \sum_{j=0}^{m-1} r_j \sqrt{\rho^j}, r_j \in F \) then \( \sigma_i(\eta_1) = \sum_{j=0}^{m-1} r_j \zeta_m^j \sqrt{\rho}. \) Since \( \tau \) is a homomorphism

\[
\tau(\sigma_i(\eta_1)) = \sum_{j=0}^{m-1} r_j \tau(\zeta_m^j) \tau(\sqrt{\rho^j}).
\]

Hence by definition of \( \tau \)

\[
\tau(\sigma_i(\eta_1)) = \sum_{j=0}^{m-1} r_j \zeta_m^{(i+2)j} \sqrt{\rho^j} = \sigma_{i+2}(\eta_1).
\]

Likewise \( \tau(\sigma_i(\eta_2)) = \sigma_{i+2}(\eta_2) \). If necessary the indices \( i \) and \( i + 2 \) have to be taken mod \( m \).

Therefore for \( j = 1, 2, \)

\[
\tau \left( \frac{1}{\rho_j} \sigma_{2(h-1)}(\eta_j) \left( \sigma_{2(h-2)}(\eta_j) \right)^{d-1} \right) = \frac{1}{\rho_j} \sigma_{2h}(\eta_j) \left( \sigma_{2(h-1)}(\eta_j) \right)^{d-1} = \tau(\eta).
\]

On the other hand, by definition of \( \rho_1, \rho_2 \) and \( \eta_1, \eta_2 \), and by the formulas for \( \tau(\sigma_i(\eta_1)), \tau(\sigma_i(\eta_2)) \),

\[
\tau(\eta) = \frac{1}{\rho_1} \sigma_{2h}(\eta_1) \left( \sigma_{2(h-1)}(\eta_1) \right)^{d-1} = \zeta^{(2h)} \sqrt[4]{\sigma_{2h} (\gamma)} \left( \zeta^{(2(h-1))} \sqrt[4]{\sigma_{2(h-1)} (\gamma)} \right)^{d-1}
\]

and

\[
\tau(\eta) = \frac{1}{\rho_2} \sigma_{2h}(\eta_2) \left( \sigma_{2(h-1)}(\eta_2) \right)^{d-1} = \zeta^{(2h)} \sqrt[4]{\sigma_{2h} (\gamma)} \left( \zeta^{(2(h-1))} \sqrt[4]{\sigma_{2(h-1)} (\gamma)} \right)^{d-1}.
\]

Therefore

\[
\zeta^{(2h)} \sqrt[4]{\sigma_{2h} (\gamma)} \left( \zeta^{(2(h-1))} \sqrt[4]{\sigma_{2(h-1)} (\gamma)} \right)^{d-1} = \zeta^{(2h)} \sqrt[4]{\sigma_{2h} (\gamma)} \left( \zeta^{(2(h-1))} \sqrt[4]{\sigma_{2(h-1)} (\gamma)} \right)^{d-1}.
\]

This implies

\[
\zeta^{(2h)} = \zeta^{(2h)}.
\]
which was to be shown.

To prove the claim also for the odd indices we can use the same proof except that we start with the fact that \( \zeta^{(1)} = \zeta'^{(1)} \) and \( \zeta^{(3)} = \zeta'^{(3)} \).

To prove the lemma for complex fields containing a primitive \( m \)-th root of unity \( \zeta_m \) note \( F = F(\zeta_m) \). Hence the mapping \( \tau \) from above may simply be chosen as the isomorphism mapping \( \sqrt[p]{\rho} \) onto \( \zeta_m \sqrt[p]{\rho} \). Accordingly, \( \tau(\sigma_i(\eta)) = \sigma_{i+1}(\eta) \) for all elements \( \eta \) in \( F(\sqrt[p]{\rho}) \) and we need not distinguish between odd and even indices.

Clearly, the proof for Lemma 6.4.3 shows that in case \( F, E \subset \mathbb{R} \) the prefix \( (\zeta^{(0)}, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)}) \) determines uniquely an admissible sequence. Accordingly, in the complex case any two admissible sequences with common prefix \((\zeta^{(0)}, \zeta^{(1)})\) are equal. But we will use the result only for normalized admissible sequences.

The proof also shows that if the degree of \( \sqrt[p]{\rho} \) over \( F(\zeta_m) \) is \( m \) then in the real case \( \zeta^{(1)} \) alone determines a normalized admissible sequence, too. In a practical implementation of the algorithms described below this may be interesting for reasons of efficiency. But for our purposes, which is basically a polynomial time algorithm, the lemma above is sufficient. Hence we will not pursue this observation in the sequel.

Observe that the proof of Lemma 6.4.3 more or less is an algorithm to compute the superset for the set of admissible sequences. As in case of Lemma 6.3.2 we explain the details in the next section when analyzing the denesting algorithms.

Since in the real case any normalized admissible sequence is determined by the elements \( \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)} \) the number of normalized admissible sequences for simple radical extensions is bounded by \( d^3 \). Combining this with Lemma 6.3.2 would lead to a denesting set of size \( md^3 \) if any triple \((\zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)})\) determined an admissible sequence. But we know already from Lemma 6.2.2 that the size of a denesting set can be bounded by \( m \). Hence most triples will not lead to a normalized admissible sequences and the question is whether we can do better than Lemma 6.4.3. For an arbitrary field \( F \) so far we can not, but if \( F \) is itself a real radical extension \( Q(\sqrt[q_1]{q_1}, \ldots, \sqrt[q_k]{q_k}) \) of \( Q \) and \( \sqrt[p]{\rho} \) is also a radical over \( Q \) we will show in the next section that we can compute efficiently a superset of the set of normalized admissible sequences of size at most \((24m)^3\). This is still far from \( m \) but unlike \( d^3 \) it is independent of \( d \) which is what we would expect.

The result is based on the following observation.
Assume that there are more than $R$ normalized admissible sequences such that any two of them differ in the $(i + 1)$-st component, where $i$ is fixed. By Lemma 6.4.1 this implies that more than $R$ different $d$-th roots of unity, say $\zeta^{(1)}, \ldots, \zeta^{(R+1)}$, exist such that

$$d\sqrt[d]{\frac{\gamma_{d-1}}{\sigma_i(\gamma)}} \in F(\sqrt{\rho}, \zeta_m), j = 1, \ldots, R + 1,$$

Hence any ratio of two of these elements is in $F(\sqrt{\rho}, \zeta_m)$. Especially, the ratio of any element with the first one is contained in $F(\sqrt{\rho}, \zeta_m)$. These ratios are

$$\frac{\zeta^{(1)}}{\zeta^{(j)}}, j = 1, \ldots, R + 1,$$

which, by assumption that the admissible sequences differ in the $(i + 1)$-st component, must be different $d$-th roots of unity. Hence $F(\sqrt{\rho}, \zeta_m)$ contains $R + 1$ different $d$-th roots of unity. Applying the following theorem which will be proven in an appendix shows that $R \leq 24m$.

**Theorem 6.4.4** Let $F = \mathbb{Q}(\sqrt[q_1]{\alpha}, \ldots, \sqrt[q_k]{\alpha})$ be a real radical extension of $\mathbb{Q}$. If $\zeta_m$ is a primitive $m$-th root of unity then $F(\zeta_m)$ contains at most $24m$ different roots of unity. Moreover, the constant $24$ is best possible, i.e., there are real radical extensions $F$ of $\mathbb{Q}$ and $m \in \mathbb{N}$ such that $F(\zeta_m)$ contains exactly $24m$ different roots of unity.

In particular, for each of the components 2, 3, and 4, which determine the whole sequence there exist only $24m$ possible $d$-th roots of unity and we can compute these roots by determining for which $d$-th roots of unity $\zeta$

$$\zeta \sqrt[d]{\frac{\gamma_{d-1}}{\sigma_i(\gamma)}},$$

is an element of $F(\sqrt{\rho}, \zeta_m)$. If we denote the roots that pass the test for $\sigma_1, \sigma_2, \sigma_2$ by $Z_1, Z_2, Z_3$, respectively, in order to compute a superset of normalized admissible sequences we have to consider only the $(24m)^3$ triples in $Z_1 \times Z_2 \times Z_3$.

We are now in a position to outline the basic denesting algorithms in case $F$ is an algebraic number field $\mathbb{Q}(\alpha)$. A more detailed description of the single steps and a careful analysis will be given in the next section.
6.5 Denesting Radicals - The Algorithms

We are now in a position to outline the basic denesting algorithms in case $F$ is an algebraic number field $\mathbb{Q}(\alpha)$. Explanations to the algorithms will be given partly in the descriptions itself (in italics) and partly at the end of the descriptions. A detailed description of the single steps and a careful analysis will be given in the next section.

Let us begin with an algorithm that computes for a nested radical $\sqrt[k]{\gamma}$ over an algebraic number field $F = \mathbb{Q}(\alpha)$ an element denesting it.

**Algorithm Denesting Element**

**Input** A nested radical $\sqrt[k]{\gamma}$, where $\gamma$ is a rational expression in the radicals $\sqrt[p_1]{\psi_1}, \sqrt[p_2]{\psi_2}, \ldots, \sqrt[p_K]{\psi_K}$, $\psi_i \in F$. $F$ is either real, the radicals $\sqrt[p_i]{\psi_i}$ are positive real radicals, and $\sqrt[k]{\gamma} \in \mathbb{R}$, or $F$ contains a primitive $k$-th root of unity and $\psi_i$ divides $d$ for all $i$.

**Output** “Yes”, and elements $\gamma_0 \in F$, $\eta \in F(\sqrt[p_1]{\psi_1}, \sqrt[p_2]{\psi_2}, \ldots, \sqrt[p_K]{\psi_K})$ such that $\sqrt[k]{\gamma_0} \sqrt[k]{\gamma} = \eta$ if some element in $F$ leads to a denesting, “No”, otherwise.

**Step 1** Denote $F(\sqrt[p_1]{\psi_1}, \ldots, \sqrt[p_i]{\psi_i})$ by $F(i)$. Determine for each $i$ the relative degree $n_i = [F(i) : F(i-1)]$ and a primitive element $\eta_i$ of $F(i)$ over $\mathbb{Q}$ represented by its minimal polynomial $p_i$.

(This is basically a preprocessing step. The elements $\eta_i$ are needed for applying the algorithm leading to Theorem 5.2.6, p. 61, to the fields $F(i)$.)

Renumber the radicals such that $n_i > 1$ for the first $k$ indices and $n_i = 1$ for the remaining indices. Hence $F(\sqrt[p_1]{\psi_1}, \sqrt[p_2]{\psi_2}, \ldots, \sqrt[p_K]{\psi_K}) = F(\sqrt[p_1]{\psi_1}, \sqrt[p_2]{\psi_2}, \ldots, \sqrt[p_K]{\psi_K})$. Also compute the set $E(i)$ of those $s \in \mathbb{N}$, $s < n_i$, such that $\sqrt[p_s]{\psi_s}$ divides $d$ for all $d$.

(See Lemma 6.2.2 and observe that $\{1, \sqrt[p_1]{\psi_1}, \ldots, \sqrt[p_{n_i-1}]{\psi_i}\}$ is a basis for $F(i)$ over $F(i-1)$.)

Using Theorem 5.2.6, p. 61, compute the representation of $\gamma$ with respect to the primitive element $\eta_k$ of $F(k)$ and determine a rational integer $C$ such that $\gamma(k) = C\gamma$ is an algebraic integer.

(The integer $C$ will prevent an exponential coefficient growth throughout the next steps of the algorithm.)
Step 2 Set $D^{(k)} := \{\gamma^{(k)}\}$.

For $i = 0, 1, \ldots, k - 1$ do the following:

Using the characterization in Lemma 6.4.3 and the algorithm leading to Theorem 5.2.6, p. 61, compute for all elements $\gamma^{(k-i)}$ in $D^{(k-i)}$ a superset $A$ of the set of normalized admissible sequences for $d\sqrt{\gamma^{(k-i)}}$ over $F^{(k-i-1)}$.

For each element $\gamma^{(k-i)}$ in $D^{(k-i)}$ do the following:

Until a denesting element for $d\sqrt{\gamma^{(k-i)}}$ over $F^{(k-i-1)}$ has been found, using the method of Theorem 5.2.6, p. 61, determine for all sequences in $A (\zeta^{(0)}, \zeta^{(1)}, \ldots, \zeta^{(n_{k-i}-1)}), \zeta^{(0)} = 1$, and for all $r \in \{0, 1, \ldots, n_{k-i} - 1\}$ the exact representation of the element in $F^{(k-i-1)}$ that

$$\left( \sum_{j=0}^{n_{k-i}-1} \zeta^{(j)} \sigma_j \left( d \sqrt{\rho_{k-i}} \right) \sqrt{\sigma_j(\gamma^{(k-i)})} \right)^d$$

has to be if it is an element of $F^{(k-i-1)}$. The representation is with respect to the primitive element $\eta_{k-i-1}$ for $F^{(k-i-1)}$. Determine whether the inverse of this element in $F^{(k-i-1)}$ denests $d\sqrt{\gamma^{(k-i)}}$ over $F^{(k-i-1)}$.

(By Lemma 6.3.2 and Remark 6.3.3 we know that if a denesting element for $d\sqrt{\gamma^{(k-i)}}$ exists then the inverse of at least one denesting element can be described by the formula above.)

If the inverse $\gamma^{(k-i-1)}$ of a denesting element has been found, set

$$D_{\gamma^{(k-i-1)}} = \{\gamma^{(k-i-1)} d s \sqrt{\rho_{k-i-1}} s \in E^{(k-i)}\}.$$

(By Lemma 6.2.2 the set $D_{\gamma^{(k-i-1)}}$ contains the inverses of an irreducible denesting set for $d\sqrt{\gamma^{(k-i)}}$ over $F^{(k-i-1)}$.)

119
Set
\[ D^{(k-i-1)} = \bigcup_{\gamma^{(k-i)} \in D^{(k-i)}} D_{\gamma^{(k-i)}}. \]

If \( D^{(k-i-1)} \) is empty, stop and output “No”, otherwise proceed with the next \( i \).

**Step 3** Using the algorithm leading to Theorem 5.2.6, p. 61, decide whether some element \( \gamma^{(0)} \) in \( D^{(0)} \) has the property
\[
\sqrt[d]{\frac{C}{\gamma^{(0)}}} = \eta \in F\left( \sqrt[d]{\rho_1}, \sqrt[d]{\rho_2}, \ldots, \sqrt[d]{\rho_k} \right),
\]
if this is the case determine the representation of \( \sqrt[d]{\frac{C}{\gamma^{(0)}}} \) as a linear combination of the elements of the standard basis, output this representation and \( \frac{C}{\gamma^{(0)}} \), otherwise output “No”.

Observe that if \( D \) is a denesting set for \( \sqrt[d]{C\gamma} \) then multiplying the elements in \( D \) by \( C \) yields a denesting set for \( \sqrt[d]{\gamma} \). Therefore the proof of correctness for this algorithm follows immediately from Lemma 6.2.4 and Lemma 6.3.2 and Remark 6.3.3. In fact, since \( \{1, a_k\sqrt[k]{\rho}, \ldots, a_k\sqrt[k]{\rho^{n_k-1}}\} \) is a basis for \( F^{(k-i)} \) over \( F^{(k-i-1)} \) it follows from Lemma 6.3.2 and Remark 6.3.3 that if a nested radical \( \sqrt[d]{\gamma^{(k-i)}} \) can be denested over \( F^{(k-i-1)} \) by a denesting element then such a denesting will be computed in **Step 2**. From Lemma 6.2.4 it follows that the inverses in \( D^{(0)} \) are a superset of a denesting set for \( \gamma^{(k)} = C\gamma \).

As mentioned in the algorithm computing a denesting set for \( \sqrt[d]{C\gamma} \) rather than for \( \sqrt[d]{\gamma} \) avoids an exponential growth in the coefficient size of the elements in \( D^{(i)} \).

The last step has been included since it is crucial for the **General Denesting Algorithm** to be described below.

The set \( D^{(k)} \) has one element and the set \( D^{(k-1)} \) has at most \( n_k \) elements which is the degree of \( F^{(k)} \) over \( F^{(k-1)} \) (see Corollary 6.2.3). By induction on \( i \) one shows that the set \( D^{(k-i)} \) has at most \( \prod_{j=0}^{i-1} n_{k-j} \) elements. Hence admissible sequences have to be computed for at most
\[
\sum_{i=0}^{k-1} |D^{(k-i)}| = N \sum_{i=1}^{k} \frac{1}{n_1 n_2 \cdots n_i} \leq N \sum_{i=1}^{k} \frac{1}{2^i} \leq N
\]
elements.
And if for each of the elements in $D^{(k-i)}$ the admissible sequence determined by the prefix $(1, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)})$ or $(1, \zeta^{(1)})$ can be computed efficiently at each level of the algorithm we have to check for at most $d^3 n_{k-i} |D^{(k-i)}|$, $i = 0, \ldots, k - 1$, or $dn_{k-i} |D^{(k-i)}|$, $i = 0, \ldots, k - 1$, different complex numbers fitting into the description of Lemma 6.3.2 whether they correspond to a denesting element in $F(\sqrt[d]{\rho_1}, \sqrt[d]{\rho_2}, \ldots, \sqrt[d]{\rho_{k-i}})$. The factor $n_{k-i}$ occurs because we have to apply the formula in Lemma 6.3.2 for all combinations of normalized admissible sequences and elements of a basis of $F^{(k-i)}$ over $F^{(k-i-1)}$.

Hence overall this step has to be done at most

$$d^3 \frac{N}{n_1 + \frac{1}{n_1} + \frac{1}{n_1 n_2} + \ldots + \frac{1}{n_1 n_2 \ldots n_{k-1}}} \leq 2d^3 N$$

times in the real case, and accordingly, at most $2dN$ times in the complex case. Since $|D^{(0)}| \leq N$ the last step has to be applied at most $N$ times.

**Remark 6.5.1** In the complex case, the algorithm can be improved. $D^{(0)}$ is a superset of a denesting set of $\sqrt[d]{\gamma}$. On the other hand, it contains at most $N$ elements. By Corollary 6.2.3 in this case any denesting set has size at least $N$. Therefore, $D^{(0)}$ must be an irreducible denesting set for $\sqrt[d]{\gamma}$. Hence if $\sqrt[d]{\gamma_0} \in F$ for some $\gamma_0 \in F$ then any inverse of an element in $D^{(0)}$ must denest $\sqrt[d]{\gamma}$. Therefore we only need to compute a single element of $D^{(0)}$. Accordingly, in Step 2 on each level we compute the inverse of a single element leading to a denesting, i.e., for all $i D^{(i)}$ consists of a single element. This observation saves us basically a factor of $N$ in the final run time. We will analyze this slightly modified algorithm in the complex case.

By Theorem 6.1.4 in the complex case the Algorithm Denesting Element finds already a denesting of $\sqrt[d]{\gamma}$ using radicals of the form $\sqrt[d]{\rho}$, $\rho \in F$, $t \in \mathbb{N}$, $t$ dividing $d$, if any such denesting exists. In the real case however, the Algorithm Denesting Element does not necessarily find a denesting using only real radicals if any such denesting exists.

The correctness of the following Algorithm Real Denesting is immediate from Theorem 6.1.1.

**Algorithm Real Denesting**

**Input** A nested radical $\sqrt[d]{\gamma}$ as in the real case of the Algorithm Denesting Element.

121
Output A denesting for $\sqrt[\gamma]{}$ using real radicals if such a denesting exists.

Step 1 Apply Step 1 from the Algorithm Denesting Element. Denote by $B = \{\beta_0, \beta_1, \ldots, \beta_{N-1}\}$ the standard basis for the radical extension $F^{(k)}$ over $F$.

Step 2 Applying Step 2 and Step 3 of the Algorithm Denesting Element check whether any nested radical $\sqrt[\beta_j\gamma]{}$ for $j = 0, 1, 2, \ldots, N - 1$, denests.

If this is not the case, output "No".

If some $\sqrt[\beta_j\gamma]{}$ can be denested by an element in $F$ output this denesting and $\beta_j$ as a denesting for $\gamma$.

By the description of Step 2 of the Algorithm Real Denesting we mean that Step 2 and Step 3 of the Algorithm Denesting Element are successively applied with $D^{(k)} = \{\beta_j\gamma^{(k)}\}$, $j = 0, 1, \ldots, N - 1$. The reader may wonder why we do not simply apply the Algorithm Denesting Element to all radicals $\sqrt[\beta_j\gamma]{}$. The answer is that Step 1 of the Algorithm Denesting Element is the same for all these radicals.
The correctness of the following General Denesting Algorithm follows directly from Theorem 6.1.2 and its complex equivalent Theorem 6.1.5.

**The General Denesting Algorithm**

**Input** $k$ depth 2 nested radicals $\sqrt[k]{\gamma_i}$ over a field $F$ and $k$ depth one radical expressions $\kappa_i$ over $F$. If $F$ is real then $\sqrt[k]{\gamma_i} \in \mathbb{R}$ and the radicals appearing in $\gamma_i$ or $\kappa_i$ are real. If $F$ is not real then it contains a primitive $d$-th root of unity, all $d_i = d$ and the depth 1 radicals appearing in $\gamma_i$ or $\kappa_i$ are of the form $\sqrt[d]{\rho}$, $\rho \in F$, $d' | d$.

**Output** In the real case, a denesting of $S = \sum_{i=1}^{k} \kappa_i \sqrt[k]{\gamma_i}$ using only real radicals if any exists, in the complex case, a denesting of $S = \sum_{i=1}^{k} \kappa_i \sqrt[k]{\gamma_i}$ using only radicals of the form $\sqrt[d]{\rho}$, $\rho \in F$, $d' | d$ if any exists.

**Step 1** Use the following procedure to partition the set of nested radicals $\{ \sqrt[k_1]{\gamma_1}, \sqrt[k_2]{\gamma_2}, \ldots, \sqrt[k_k]{\gamma_k} \}$ into subsets $R_t$ such that two nested radicals are in the same subset if and only if their ratio can be written either as a sum of real depth 1 radicals over $F$ in the real case, or as a sum of depth 1 radicals of the form $\sqrt[d]{\rho}$, $\rho \in F$, for some $d' | d$, in the complex case.

In the complex case, apply the Algorithm Denesting Element to $\sqrt[\ell]{\gamma_t} \sqrt[^{\ell-1}]{\gamma_t} = \gamma_t \sqrt[\ell]{\gamma_t}$ for all pairs $\gamma_i, \gamma_t$ of radicals. If a radical denests denote the denesting by $\gamma_t$.

In the real case, for any pair of radicals determine whether $\sqrt[^{\ell-1}]{\gamma_t} \gamma_t = \gamma_t \sqrt[^{\ell-1}]{\gamma_t}$ denests using real radicals by applying the Algorithm Real Denesting. Again denote a denesting by $\gamma_t$.

**Step 2** Assume $\sqrt[k_t]{\gamma_t} \in R_t$, for $t = 1, 2, \ldots, h$, where $h$ is the number of subsets in the partition of **Step 1**.

Write

$$S = \sum_{t=1}^{h} \gamma_t \sum_{\gamma_i \in R_t} \kappa_i \gamma_i.$$

Determine whether any $\sqrt[k_t]{\gamma_t}$, $t = 1, 2, \ldots, h$, denests. If so, let $\sqrt[k_t]{\gamma_t}$ be the one that denests. (Observe that only one nested radical $\sqrt[k_t]{\gamma_t}$, $t = 1, \ldots, h$, can denest. If two of them did, so would their ratio, which is impossible because of the partition determined in **Step 1**.)
Step 3 Transform each expression $\kappa_i$ into a sum of radicals.

If $\sqrt[\gamma_1]{\gamma}$ denests, check whether

$$\sum_{\{i \in \mathbb{N} | \sqrt[\gamma_1]{\gamma_i} \in R_i\}} \kappa_i \gamma_i = 0$$

for all $t \geq 2$. If so output $\sqrt[\gamma_1]{\gamma} \sum_{\{i \in \mathbb{N} | \sqrt[\gamma_1]{\gamma_i} \in R_i\}} \kappa_i \gamma_i$ as the denesting for $S$. If at least one sum is not zero output that $S$ cannot be denested.

If $\sqrt[\gamma_1]{\gamma}$ does not denest, check whether all sums are zero and, if so, output zero as a denesting for $S$, otherwise output that it cannot be denested.

The correctness of this algorithm follows directly from Theorem 6.1.2 and its complex equivalent Theorem 6.1.5.

Observe that by the Algorithm Denesting Element when we have to check whether $\sum_{\{i \in \mathbb{N} | \sqrt[\gamma_1]{\gamma_i} \in R_i\}} \kappa_i \gamma_{ij}$ is zero this sum is a sum of radicals. $\gamma_{it}$ is a sum of radicals by the last step in the Algorithm Denesting Element and for $\kappa_i$ a representation as a sum of radicals has been determined in the General Denesting Algorithm itself. Hence we can apply the algorithms of Section 5 to determine whether the sums $\sum_{\{i \in \mathbb{N} | \sqrt[\gamma_1]{\gamma_i} \in R_i\}} \kappa_i \gamma_{it}$ are zero.
7 Denesting Radicals - The Analysis

In this section we analyze the algorithms Algorithm Denesting Element, Algorithm Real Denesting, and the General Denesting Algorithm. We will show how to fill in the details into the descriptions of the previous section such that the Algorithm Denesting Element and the Algorithm Real Denesting run in time polynomial in \( d \), the degree \( N \) of the radical extension generated by the radicals appearing in the description of \( \gamma \), and in the input size of the problem. For the General Denesting Algorithm we will show that it runs in time polynomial in the \( d_i \)’s, in \( N \), which is the maximum degree of an extension generated by the radicals appearing in a single pair \( \gamma_i, \kappa_i \) of radical expressions, and in the input size of the problem.

Recall from Section 6 that for the Algorithm Denesting Element and the Algorithm Real Denesting the output, which will be a sum of depth 1 radicals, may have \( N \) terms. Moreover, if \( \gamma \) is a sum of radicals, \( N \) will be the degree of the minimal polynomial of \( \gamma \) (see Theorem 3.11, p. 31, or Theorem 3.13, p. 32). Hence in this case we achieve a denesting algorithm whose run time is polynomial in the description size of the minimal polynomial of the nested radical. This is the first algorithm to achieve such a run time.

As will be seen later for the General Denesting Algorithm the output may have \( O(kN^2) \) terms, where \( k \) is the number of nested radicals \( \sqrt[\gamma_i]{} \) and \( N \) is as above. Moreover, the degree of the minimal polynomial of \( S \) may be \( \Omega(N^k \prod d_i) \). On the other hand, the General Denesting Algorithm runs in time polynomial in \( k, \max\{d_i\} \), and \( N \). Again this is the first algorithm to achieve such a run time.

As is clear from the previous section we basically have to analyze the Algorithm Denesting Element.

7.1 Preliminaries

In this subsection we deduce several results that will simplify the analysis of the three main steps in the Algorithm Denesting Element. The main part of the analysis of the Algorithm Denesting Element will be done in the following three subsections. In the final subsection the results will be combined to analyze the overall run times of the three denesting algorithms.

As in Section 6 we refer to the two different types of denesting problems we consider as the real and complex case.

As in Section 5 we assume that an algebraic number field \( F = \mathbb{Q}(\alpha) \) is generated by an algebraic integer \( \alpha \). The field is specified by the minimal
polynomial \( p(X) = \sum_{i=0}^{n} p_i X^i \), \( p_i \in \mathbb{Z} \), \( p_n = 1 \), of \( \alpha \). Throughout this section \( n \) always denotes the degree of \( p \) and we assume that the length of \( p \) satisfies \( |p|_2 < 2^l \). Furthermore it is assumed that \( \alpha \) is distinguished from its conjugates by an isolating interval or rectangle, that is, an interval or rectangle containing no root of \( p \) except \( \alpha \). In general, we will not mention this interval.

Any element \( \beta \) of the algebraic number field \( \mathbb{Q}(\alpha) \) is specified by an \( (n+1) \)-tuple \( (b_0, b_1, \ldots, b_n) \in \mathbb{Z}^{n+1} \), \( \gcd(b_0, b_1, \ldots, b_n) = 1 \), such that \( \beta = \frac{1}{d} \sum_{i=0}^{n-1} b_i \alpha^i \). For the definition of the infinity norm \( |\beta|_\infty \) and and the coefficient size \( |\beta| \) we refer to Section 5.1.

A radical \( \sqrt[d]{\rho}, d \in \mathbb{N}, \rho \in \mathbb{Q}(\alpha) \), is represented only by \( d \) and \( \rho \). To avoid ambiguity in this section \( \sqrt[d]{\rho} \) always has the value

\[
\sqrt[d]{\rho} = |\rho|^{\frac{1}{d}} \left( \cos \left( \frac{1}{d} \phi \right) + i \sin \left( \frac{1}{d} \phi \right) \right),
\]

where \( \phi \in (-\pi, \pi) \) is the angle of \( \rho \) if written in polar coordinates, and \( |\rho|^{\frac{1}{d}} \) is the positive real root of \( |\rho| \). Recall from Section 6 that this is no restriction\(^{15}\).

Recall that we can check whether an algebraic number field is real or contains a certain primitive root of unity by algorithms whose run times are polynomial in the description size of the minimal polynomial \( p \) and, in case of the root of unity, in the degree of the root of unity (see for example Theorem 5.6.5, p. 92). Likewise, if \( \mathbb{Q}(\alpha) \subset \mathbb{R} \), we can check whether an element \( \beta \) in \( \mathbb{Q}(\alpha) \) is positive in time polynomial in the description size of \( p \) and \( \beta \). Hence for the Algorithm Denesting Element, as described in the previous section, we can check the input conditions on \( \mathbb{Q}(\alpha) \) or on the radicals \( \sqrt[d]{\rho} \) efficiently.

We want to apply the Algorithm Denesting Element or the Algorithm Real Denesting to a nested radical \( \sqrt[d]{\gamma} \), where \( \gamma \) is a rational expression in radicals over \( \mathbb{Q}(\alpha) \). That is, \( \gamma \) is an expression built up from elements in \( \mathbb{Q}(\alpha) \) and from elements in a finite set \( \{ \sqrt[d_1]{\rho_1}, \sqrt[d_2]{\rho_2}, \ldots, \sqrt[d_K]{\rho_K} \} \) of radicals over \( \mathbb{Q}(\alpha) \) using the arithmetic operations addition, subtraction, multiplication, and division.

We may for example assume that \( \gamma \) is given as a straight-line program using elements in \( \mathbb{Q}(\alpha) \) and the radicals \( \sqrt[d]{\rho} \). As is not hard to see analyzing

\(^{15}\)We may also use the same convention as in Section 5, but in view of the definitions in Section 6 (in particular, admissible sequences) choosing this more restricted form is appropriate.
the Algorithm Denesting Element for γ’s defined in this way would lead to an algorithm whose run time is exponential in the number of steps of the straight-line program. Instead of assuming a particular input form below we state five conditions on the representation of γ and the analysis we give in the sequel applies to all classes of expressions satisfying these conditions. By an appropriate choice of the parameters also nested radicals defined via straight-line programs fulfill these conditions. The conditions are as follows.

1) $\sqrt[d]{\rho}$ is an algebraic integer for all $i = 1, 2, \ldots, K$.
2) $\sqrt[d]{\rho}$ is of degree $d_i$ over $F = \mathbb{Q}(\alpha)$.
3) The radicals $\sqrt[d]{\rho_i}$ are linearly independent over $\mathbb{Q}(\alpha)$.
4) An integer $K \in \mathbb{N}$ is known such that $\lceil \gamma \rceil < 2^K$ and such that an integer $C \in \mathbb{N}$, $C < 2^K$ exists with $C\gamma$ is an algebraic integer.
5) For any positive $\epsilon$ an approximation to $\gamma$ with absolute error less than $\epsilon$ can be determined by elementary operations on floating-point numbers of size polynomial in $\log \frac{1}{\epsilon}$ and $K$. The number of operations necessary is also bounded by a polynomial in $\log \frac{1}{\epsilon}$ and $K$.

In 4) we do not assume that we know the integer $C$. Rather it will be determined in the algorithm.

The first three conditions are of a more technical nature, they simplify the notation and reduce the number of input parameters but are not crucial to the analysis. We will show that these conditions can always be satisfied.

Conditions 4) and 5) are the basic assumptions of the analysis. In fact, $K$ will be one of the important run time parameters. Basically, using the methods of Section 5), in particular by Theorem 5.2.6, p. 61, Conditions 4) and 5) allow us to determine some canonical basis representation of $\gamma$. The denesting algorithm will work with this canonical representation rather than with the original expression for $\gamma$. Moreover, these conditions will allow us to determine whether $\gamma$ is positive in case $F \subset \mathbb{R}$ and $d$ is even.

To justify and demonstrate the generality of the conditions stated above we will show below that a large class of expressions satisfy, in particular,
conditions 4) and 5). Hence the analysis applies to these expressions. But due to the generality of 4) and 5) the analysis is not restricted to expressions of the form described below.

Let us begin by considering the first three conditions.

Let \( r_i \) be the denominator of \( \rho_i \). As shown in the proof of Lemma 5.3.1, p. 63, \( r_i \sqrt{\rho_i} \) is an algebraic integer. This shows how to satisfy condition 1) by replacing \( \sqrt{\rho_i} \) by \( r_i \sqrt{\rho_i} \) and from now on we assume that all elements \( \sqrt{\rho_i} \) are integers.

Condition 2) can be satisfied by determining for any element in \( \{ \sqrt{\rho_1}, \sqrt{\rho_2}, \ldots, \sqrt{\rho_K} \} \) the smallest \( m_i \) such that \( d_i \sqrt{\rho_i} \) is an algebraic integer. This shows how to satisfy condition 1) by replacing \( d_i \sqrt{\rho_i} \) by \( r_i d_i \sqrt{\rho_i} \) and from now on we assume that all elements \( \sqrt{\rho_i} \) are integers.

Condition 2) can be satisfied by determining for any element in \( \{ \sqrt{\rho_1}, \sqrt{\rho_2}, \ldots, \sqrt{\rho_K} \} \) the smallest \( m_i \) such that \( d_i \sqrt{\rho_i} \) is an algebraic integer. This shows how to satisfy condition 1) by replacing \( d_i \sqrt{\rho_i} \) by \( r_i d_i \sqrt{\rho_i} \) and from now on we assume that all elements \( \sqrt{\rho_i} \) are integers.

By Theorem 3.3, p. 24, this is the degree of \( d_i \sqrt{\rho_i} \) over \( \mathbb{Q}(\alpha) \). Computing the integers \( m_i \) can easily be done by checking successively for \( m_i = 1, 2, \ldots \) whether \( d_i \sqrt{\rho_i} \) is an algebraic integer. This shows how to satisfy condition 1) by replacing \( d_i \sqrt{\rho_i} \) by \( r_i d_i \sqrt{\rho_i} \) and from now on we assume that all elements \( \sqrt{\rho_i} \) are integers.

Let \( S_j \) denote the set \( \{ \sqrt{\rho_1}, \sqrt{\rho_2}, \ldots, \sqrt{\rho_K} \} \) for \( \gamma \) by applying the algorithm leading to Theorem 5.6.1, p. 88, or Theorem 5.6.2, p. 90, to the set \( \{ \sqrt{\rho_1}, \sqrt{\rho_2}, \ldots, \sqrt{\rho_K} \} \) in order to determine a set of linearly independent radicals of maximum size. Recall that any other radical can be written as a product of an element in this set and an element in \( \mathbb{Q}(\alpha) \) (Corollary 3.10, p. 28). Moreover the multiples in \( \mathbb{Q}(\alpha) \) will have coefficient size less than \( 2^{3L} \) (Lemma 5.3.2, p. 64).

Finally, Condition 3) can be satisfied for any expression \( \gamma \) by applying the algorithm leading to Theorem 5.6.1, p. 88, or Theorem 5.6.2, p. 90, to the set \( \{ \sqrt{\rho_1}, \sqrt{\rho_2}, \ldots, \sqrt{\rho_K} \} \) in order to determine a set of linearly independent radicals of maximum size. Recall that any other radical can be written as a product of an element in this set and an element in \( \mathbb{Q}(\alpha) \) (Corollary 3.10, p. 28). Moreover the multiples in \( \mathbb{Q}(\alpha) \) will have coefficient size less than \( 2^{3L} \) (Lemma 5.3.2, p. 64).

Let \( S_j \), \( j = 1, 2, \ldots, m \), have the form

\[
S_j = \sum_{i=1}^{m} \frac{P_{i,j}}{Q_{i,j}},
\]
where \( P_{i,j}, Q_{i,j}, i=1, \ldots, m \), are linear combinations of the radicals \( \sqrt[n]{\rho_h} \) with the \( \rho_h \)'s and the coefficients \( \kappa \) in \( \mathbb{Q}(\alpha) \) satisfying \( [\rho_h], [\kappa] < 2^L \). Furthermore assume \( \gamma \) has the form

\[
\gamma = P(S_1, S_2, \ldots, S_m),
\]

where \( P(X_1, X_2, \ldots, X_m) \) is a polynomial in \( m \) variables with coefficients in \( \mathbb{Q}(\alpha) \). Assume that \( P \) has at most \( m \) terms and that the degree is also bounded by \( m^{16} \). Let the coefficients of \( P \) have coefficient size less than \( 2^L \).

Finally we assume that conditions 1), 2), and 3) are already satisfied.

To derive the bound \( K \) in 4) consider a single ratio \( P_{i,j}/Q_{i,j} \). Multiplying \( P_{i,j} \) or \( Q_{i,j} \) with the product \( b_{i,j}, c_{i,j} \), respectively, of the denominators of its coefficients yields algebraic integers since the radicals \( \sqrt[n]{\rho} \) are already integers. We obtain

\[
|b_{i,j}|, |c_{i,j}| < 2^{NL}, \tag{1}
\]

since the number of radicals and therefore the number of terms in each \( P_{i,j}, Q_{i,j} \) is bounded by \( N \) (Condition 2).

We need a bound on a rational integer \( c'_{i,j} \) such that \( c'_{i,j} \sqrt[n]{Q_{i,j}} \) is an algebraic integer. By the previous argument a positive integer \( c_{i,j} < 2^{NL} \) exists such that \( Q'_{i,j} = c_{i,j}Q_{i,j} \) is an integer. It suffices to give a bound on the size of a rational integer \( c'_{i,j} \) such that \( c'_{i,j} \sqrt[n]{Q'_{i,j}} \) is an algebraic integer.

We need to bound the infinity norm of \( Q'_{i,j} \).

The coefficients in \( Q_{i,j}, P_{ij} \) have coefficient size bounded by \( 2^L \) and the same is true for the \( \rho_i \)'s in \( \sqrt[n]{\rho} \). Hence by Lemma 5.1.11, p. 53, the coefficients have infinity norm less than \( 2^{m+L+2} \leq 2^L \) and \( \sqrt[n]{\rho_i} \) too\(^{17} \). Therefore

\[
[P_{ij}]_{\infty}, [Q_{i,j}]_{\infty} < 2^{2L+\log N} < 2^{2NL}. \tag{2}
\]

Together with the bound on \( |c_{i,j}| \) in (1) this implies

\[
[Q'_{i,j}]_{\infty} < 2^{3NL}. \tag{3}
\]

We need the following well-known lemma.

---

\(^{16}\)The argument we give is easily extended if we want to distinguish between the number of variables, the degree of \( P \), and the number of terms of \( P \). In this case \( m \) is the maximum of the parameters above.

\(^{17}\)Going from the coefficient size to the infinity norm usually means replacing \( 2^L \) by \( 2^L \).
Lemma 7.1.1 Let $\rho \in \mathbb{C}$ be an algebraic integer of degree at most $m$ such that $[\rho]_\infty < 2^B$. Then the minimal polynomial $r$ of $\rho$ satisfies

$$|r|_\infty < 2^{mB}, \ |r|_2 < 2^{m(B+1)},$$

provided $B > \log m$.

Proof: Let $e_i(x_1, \ldots, x_h)$ denote the $i$-th symmetric function on $h$ symbols $x_1, \ldots, x_h$, i.e.,

$$e_i(x_1, \ldots, x_h) = \sum_{\{k_1, \ldots, k_{h-i}\} \subseteq \{1, \ldots, h\}} x_{k_1} \cdots x_{k_{h-i}}, \ i = 0, 1, \ldots, h - 1.$$

As is well-known for $h$ elements $x_1, \ldots, x_h \in \mathbb{C}$ the number $e_i(x_1, \ldots, x_h)$ is the coefficient $r_i$ in

$$\prod_{i=1}^{h} (X - x_i) = \sum_{i=0}^{h} r_i X^i.$$

Hence for an algebraic integer $\rho$ of degree $h$ the symmetric functions on its conjugates describe the coefficients of its minimal polynomial.

Therefore, if $\rho$ is of degree $h \leq m$ and $[\rho]_\infty < 2^B$, $B > \log m$, then the $i$-th coefficient $r_i$ in $r$ satisfies

$$|r_i| \leq \binom{h}{i} [\rho]_\infty^{h-i} < h^i [\rho]_\infty^{h-i} < 2^{hB} \leq 2^{mB}.$$ 

Equivalently, the height $|r|_\infty$ is strictly less than $2^{mB}$. Moreover,

$$|r|_2 \leq |r|_1 \leq \sum_{i=0}^{h} \binom{h}{i} [\rho]_\infty^{h-i} = (1 + [\rho]_\infty)^h < 2^{h(B+1)} \leq 2^{m(B+1)}.$$

Observe that the degree of $Q'_{i,j}$ over $\mathbb{Q}$ can be at most $nN$. Together with the bound given above for the infinity norm of $Q'_{i,j}$ this yields a bound on the height of the minimal polynomial of $Q'_{i,j}$ of

$$2^{4nN^2L}.$$
Next recall that if $f(X) = \sum_{i=1}^{h} f_i X^i$ is the minimal polynomial of an algebraic number $\rho$ then $\sum_{i=1}^{h} f_{h-i} X^i$ is the minimal polynomial of $\rho^{-1}$. Also recall that $f_i \rho$ is an algebraic integer \footnote{This argument was already used at the end of Section 2 to show that any number field can be generated by an integer.}.

Therefore an integer $c'_{i,j}$ exists whose size is bounded by $2^{4nN^2L}$ such that $c'_{i,j} \frac{1}{Q'_{i,j}}$ is an algebraic integer. As mentioned before this implies that for $Q_{i,j}$ a positive integer $c'_{i,j}$ satisfying

\begin{equation}
    c'_{i,j} < 2^{4nN^2L}
\end{equation}

exists such that $c'_{i,j} \frac{1}{Q'_{i,j}}$ is an algebraic integer.

The product over $i$ of all $c'_{i,j}$'s is an integer $C_j$ such that $C_j S_j = C_j \sum_{i=1}^{m} P_{i,j}/Q_{i,j}$ is an algebraic integer. By the bounds in (1), (4)

\begin{equation}
    |C_j| < 2^{5mnN^2L}.
\end{equation}

Since the polynomial $P$ has degree at most $m$, each $S_j$ can occur at most with exponent $m$. Raising all $C_j$'s to the $m$-th power and multiplying these powers yields an integer $C$ such that any product of $S_j$'s occurring in $\gamma$ if multiplied with $C$ yields an algebraic integer. Hence a positive integer $C$ with

\begin{equation}
    C < 2^{5m^3nN^2L}
\end{equation}

exists such that $C\gamma$ is an algebraic integer.

Next let us bound the infinity norm of $\gamma$. Recall that $Q'_{i,j}$ is a root of a polynomial $f$ whose height $|f|_{\infty}$ is bounded by $2^{4nN^2L}$. By Cauchy’s bound (Lemma 4.2.1, p. 41)

\begin{equation}
    \left[ \frac{1}{Q'_{i,j}} \right]_{\infty} < 1 + 2^{4nN^2L}.
\end{equation}

Since $\frac{1}{Q_{i,j}} = c_{i,j} \frac{1}{Q'_{i,j}}$ and $1 \leq |c_{i,j}|$ we get

\begin{equation}
    \left[ \frac{1}{Q_{i,j}} \right]_{\infty} < 2^{5nN^2L}.
\end{equation}
By Lemma 5.1.7, p. 51, for each sum $\sum_{i=1}^{m} P_{i,j}/Q_{i,j}$ its infinity norm is bounded by

$$\left( \sum_{i=1}^{m} |P_{i,j}| \right)_{\infty} \left( \frac{1}{Q_{i,j}} \right)_{\infty}.$$  \hspace{1cm} (9)

Since $[P_{i,j}]_{\infty} < 2^{2(nL+2)+\log N} < 2^{2NL}$ (see (2)) this shows

$$[S_j]_{\infty} < 2^{7n^2L+\log m}.$$  \hspace{1cm} (10)

By definition of $m$ and the bounds on the coefficients of $P$

$$[\gamma]_{\infty} < 2^{8m^2n^2L}.$$  \hspace{1cm} (11)

Hence $K$ satisfying both conditions in 4) may be chosen as

$$K = 8m^3n^2L.$$  \hspace{1cm} (12)

Finally let us consider 5). Using very rough estimates we get the following lemma that is already sufficient.

**Lemma 7.1.2** Let $\epsilon > 0$. An approximation with absolute error less than $\epsilon$ to a radical expression $\gamma$ as described above can be computed using $\mathcal{O}(n)$ elementary operations on floating-point numbers of size $\mathcal{O}(n \log N + mn^2N^2\mathcal{L})$ and $\mathcal{O}(N(\log \log N + \log N + \log L + mn^2n))$ elementary operations on floating-point numbers of size $\mathcal{O}(\log N + mn^2N^2\mathcal{L})$.

**Proof:** We claim that approximations to $\alpha$ and the radicals $\sqrt[n]{P_i}$, $i = 1, 2, \ldots, K$, with absolute error less than

$$\epsilon 2^{-31mnN^2\mathcal{L}}$$

yield an approximation to $\gamma$ as required.

In fact, if $\alpha$ is approximated with absolute error less than $\epsilon 2^{-31mnN^2\mathcal{L}}$ then the coefficients in $P$ and in the $P'_{i,j}$'s, $Q'_{i,j}$'s are approximated with absolute error less than

$$\epsilon 2^{-31mnN^2\mathcal{L}} 2^{2(l+1)n+L} < \epsilon 2^{-30mnN^2\mathcal{L}},$$

using Lemma 5.4.3, p. 69, and $N \geq 2, |\alpha| < 2^l$ (see Landau’s bound on the measure of a polynomial Lemma 5.1.2, p. 48).
The latter estimate also covers the error we may make by computing the inverses of the denominators of the coefficients in \( P, P_{i,j}, Q_{i,j} \) only with error less than \( \epsilon 2^{-31mnL} \).

These approximations to the coefficients can be determined from the initial approximations using \( \mathcal{O}(nm^2N) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon} + nmN^2L) \), since the overall number of coefficients is bounded by \( \mathcal{O}(m^2N) \) and each coefficient is the linear combination of \( \{1, \alpha, \ldots, \alpha^{n-1}\} \).

Since the coefficients in the sums \( P_{i,j}, Q_{i,j} \) and the radicals \( \sqrt[\epsilon]{\rho_i} \) are bounded in absolute value by \( 2L \) (Lemma 5.1.11, p. 53) the approximations to the coefficients together with the approximations to the radicals \( \sqrt[\epsilon]{\rho_i} \) lead to approximations \( P_{i,j}, Q_{i,j} \) with absolute error less than \( \epsilon 2^{-29mnN^2L} \). Only \( \mathcal{O}(m^2N) \) operations are needed to compute these approximations.

Next we compute approximations \( \frac{1}{Q_{i,j}}' \) to \( \frac{1}{Q_{i,j}} \) with absolute error less than \( \epsilon 2^{-18mnN^2L} \). Hence

\[
\left| \frac{1}{Q_{i,j}} - \frac{1}{Q_{i,j}}' \right| \leq \left| \frac{1}{Q_{i,j}} - \frac{1}{Q_{i,j}} \right| + \left| \frac{1}{Q_{i,j}} - \frac{1}{Q_{i,j}}' \right| \leq \epsilon 2^{-18mnN^2L} + \epsilon 2^{-18mnN^2L} < \epsilon 2^{-17mnN^2L}.
\]

The first bound follows from the estimate on \( \|1/Q_{i,j}\|_\infty \) given in (8) and Lemma 5.4.3, p. 69, again.

For all \( m^2 \) pairs \( i, j \) this step is done by \( \mathcal{O}(m^2) \) operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon} + nmN^2L) \) (see Theorem 5.4.2, p. 68).

Then the approximations for the \( P_{i,j} \)'s and \( \frac{1}{Q_{i,j}} \)'s are multiplied. By our bounds on \( \|P_{i,j}\|_\infty \) in (2) and on \( \left|\frac{1}{Q_{i,j}}\right|_\infty \) in (8) this yields approximations to the quotients \( S_j \) with absolute error less than \( \epsilon 2^{-16mnN^2L} \). The run time for this step is covered by the previous ones.

Finally we need to approximate the power products of the \( S_j \) appearing in \( P \) by determining the corresponding power products of the approximations to the \( S_j \). Then we multiply these power products with the approximations to the coefficients of \( P \), and sum up the results. For the power products we need \( \mathcal{O}(m^2) \) operations and for the remaining step \( \mathcal{O}(m) \) elementary operations, both on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon} + nmN^2L) \).

By the bound \( \|S_j\|_\infty < 2^{7nN^2L} \) (see (10)) our approximation to \( S_j \) is clearly bounded in absolute value by \( 2^{7nN^2L+1} \). Lemma 5.4.3, p. 69, applied
Once more shows that the approximations to $S_j$ lead to approximations to the power products of the $S_j$ with absolute error less than
\[ \epsilon 2^{-2mnN^2L + 4m} < \epsilon 2^{-mnN^2L}. \]

Again using the bound on $S_j$ and the bound on the coefficients of $P$ shows that the final result of our computations approximates $\gamma$ with absolute error less than $\epsilon$ as desired.

To compute the approximation to $\alpha$ and $\sqrt[d]{\rho_i}$, $i = 1, 2, \ldots, K$ requires $O(n)$ elementary operations on floating-point numbers of size $O(n \log \frac{1}{\epsilon} + mn^2N^2L)$ and $O(N \log \log \frac{1}{\epsilon} + \log(nmN^2L))$ elementary operations on floating-point numbers of size $O(\log \frac{1}{\epsilon} + nmN^2L)$ which follows from from Theorem 5.4.1, p. 68, Lemma 5.4.6, p. 71, or Lemma 5.4.9, p. 74. This proves the lemma.

This finally shows that conditions 4) and 5) are satisfied for expressions as the ones described above.

One of the results to be proven below is that any expression satisfying the conditions 1) to 5) can be transformed efficiently into a sum of linearly independent radicals. But assuming this input form from the beginning seems to be too restrictive and is not appropriate if we apply the Algorithm Denesting Element in the General Denesting Algorithm. Moreover, the transformation step requires more or less all techniques used in the denesting algorithms.

Let us summarize and repeat all our input assumptions before beginning with the analysis.

**Input assumptions for the Algorithm Denesting Element**

The input to the Algorithm Denesting Element consists of a field $F = \mathbb{Q}(\alpha)$, where $\alpha$ is an algebraic integer. The minimal polynomial $p$ of $\alpha$ has degree $n$ and length $|p|_2$ bounded by $2^d$. We assume that $F$ is either real or contains a primitive $d$-th root of unity. It furthermore consists of a nested radical $\sqrt[d]{\gamma}$, where $d$ is a positive rational integer and $\gamma$ is a radical expression in the linearly independent radicals $\sqrt[p_1]{\rho_1}, \sqrt[p_2]{\rho_2}, \ldots, \sqrt[p_K]{\rho_K}$ over $\mathbb{Q}(\alpha)$. It is also assumed that $\rho_i$ is an algebraic integer and $|\rho_i| > 2^L$. Moreover, $\sqrt[p_i]{\rho_i}$ is of degree $d_i$ over $\mathbb{Q}(\alpha)$. If $F$ is real then the radicals $\sqrt[p_i]{\rho_i}$ and $\gamma$ are real. In the complex case $d_i|d$ for all $i$. A positive integer $K$ is given and it is guaranteed that $|\gamma|_\infty < 2^K$ and that a positive integer $C$ less than $2^K$ exists such that $C\gamma$ is an algebraic integer. $N$ is the degree of
the radical extension generated by the radicals in $\gamma$. It is assumed that $L = \lceil n \log n + nL \rceil > \log N$.

Finally, we assume that for any $\epsilon > 0$ $\gamma$ can be approximated with absolute error less than $\epsilon$ in time polynomial in $\log \frac{1}{\epsilon}$ and $K$.

The assumption $L > \log N$ simplifies the analysis considerably and we adopt it only for the sake of simplicity.

The next three subsections contain the detailed description and the analysis of the three steps of the Algorithm Denesting Element. The reader should recall the steps of the Algorithm Denesting Element before reading the corresponding analysis.

7.2 Description and Analysis of Step 1

We have to show how to compute the degrees $n_i$ of the extensions $F^{(i)} = F(\sqrt[di]{\rho_1}, \ldots, \sqrt[di]{\rho_i})$ over $F^{(i-1)} = F(\sqrt[di]{\rho_1}, \ldots, \sqrt[di-1]{\rho_{i-1}})$, $i = 1, 2, \ldots, K$, and how to compute primitive elements $\eta_i$ for the extensions $F^{(i)}$ over $\mathbb{Q}$. We need these elements for efficient computations in $F^{(i)}$. We use two different representations for the elements $\eta_i$. One is via a rational integer $c$ such that $\eta_i = c\alpha + \sqrt[di]{\rho_1} + \sqrt[di]{\rho_2} + \cdots + \sqrt[di]{\rho_i}$ for all $i$. The other representation for $\eta_i$ is via its minimal polynomial $p_i$ which is computed using Schönhage’s algorithm. Finally in the real case sets $E^{(i)}$ have to be computed such that $\sqrt[di]{\rho_i} s \in F^{(i-1)}$ if and only if $s \in E^{(i)}$.

Lemma 7.2.1 In both, the real and complex case, the degrees $n_i$, primitive elements $\eta_i$, their minimal polynomials $p_i$, and the sets $E^{(i)}$ can be computed using $O(n^4 L \log N)$ elementary operations on integers of size $O(n^2 \log(NL))$ and $O(n^3 N^2 \log(NL))$ elementary operations on floating-point numbers of size $O(n^2 \log(NL))$.

The minimal polynomial $p_i$ of $\eta_i$ satisfies $|p_i| < 2^{4n^2 L}$.

Proof: We give the proof only for the real case, except for some minor changes the proof in the complex case is the same.

First we show how to determine the degrees $n_i$. We do this in the same way as in the considerations leading to Theorem 4.2.2, p. 42.

Assume that the degrees $n_j$, $j < i$, have already been computed. Then we also know the standard basis of $F^{(i-1)}$. By Lemma 3.9, p. 28, $n_i$ is the smallest integer such that $\sqrt[di]{\rho_i} \in F^{(i-1)}$. Moreover, in this case $\sqrt[di]{\rho_i}$ can be written as a product of an element in $F$ and an element of the standard basis of $F^{(i-1)}$. 135
The degree of $F^{(i-1)}$ is given by $N_{i-1} = \prod_{j=1}^{i-1} n_j$. By assumption, $d_j$ is the degree of $\sqrt[n]{\rho_j}$ over $F$. Hence $\sqrt[n]{\rho_j}$, $j \leq i-1$, generates a subfield of $F^{(i-1)}$ of degree $d_j$. Therefore $d_j$, $j \leq i-1$, must divide $N_{i-1}$. This implies that if $\ell_j := N_{i-1}/d_j$ then any element $\beta_h$ in the standard basis of $F^{(i-1)}$ can be written as

$$\beta_h = N_{i-1} \prod_{j=1}^{i-1} \rho_j^{e_j \ell_j}, e_j < n_j.$$ 

where $e_j < n_j$ hence $e_j \ell_j \leq N_{i-1}$. By Lemma 5.5.5, p. 83,

$$\left[ \prod_{j=1}^{i-1} \rho_j^{e_j \ell_j} \right] < 2 \mathcal{L} N \log N,$$

since overall at most $\log N$ degrees $n_j$ can be larger than 1, and $N$, although so far unknown, is clearly an upper bound for $N_{i-1}$. Together this implies

$$\sum_{j=1}^{i-1} e_j \ell_j \leq N \log N.$$

Using $O(n^2 N \log N)$ elementary operations on integers of size $O(\mathcal{L} N \log N)$ we can easily determine the representations as linear combinations of $\{1, \alpha, \ldots, \alpha^{n-1}\}$ of $\prod_{j=1}^{i-1} \rho_j^{e_j \ell_j}$ for all elements $\beta_h = N_{i-1} \sqrt[n]{\prod_{j=1}^{i-1} \rho_j^{e_j \ell_j}}$ in the standard basis of $F^{(i-1)}$ (Lemma 5.5.2, p. 82).

Also by Lemma 5.5.5, p. 83,

$$[\sqrt[n]{\rho_i^m}] < 2 N \mathcal{L}, \text{ if } m \leq d_i.$$

By the main result in Section 5 (see Table 1, p. 88) we can determine for each basis element $\beta_h$ and each power $\sqrt[n]{\rho_i^m}$ whether their ratio is an element of $\mathbb{Q}(\alpha)$ using $O(n)$ elementary operations on floating-point numbers of size $O(\mathcal{L} n^3 N \log N)$, $O(\log(n \mathcal{L}))$ elementary operations on floating-point numbers of size $O(\mathcal{L} n^2 N \log N)$, $O(\mathcal{L} n^4 N \log N)$ elementary operations on integers of size $O(\mathcal{L} n^2 N \log N)$, and $O(n^2 \log N)$ elementary operations on integers of size $O(\mathcal{L} n N^2 \log N)$.\(^{19}\)

Since the standard basis for $F^{(i-1)}$ contains $N_{i-1}$ elements and the ratio tests have to be applied only to $\sqrt[n]{\rho_i}, \ldots, \sqrt[n]{\rho_{i-1}}$ for each $i$ this test has to be applied at most $n_i N_{i-1}$ times. Since $n_i N_{i-1} \leq N$ and $K \leq N$ throughout the first step of the Algorithm Denesting Element the test has to be applied at most $N^2$ times. Hence within the time bounds stated in the lemma the

\(^{19}\)By applying the results from Section 5 to this case be careful about the meaning of $\mathcal{L}$ and $L$. In the present case the coefficient size of the input elements is already $2^{\mathcal{L} N \log N}$. 136
degrees $n_i$ can be determined and from now on we assume that $N_i$ and $N$ are known.

The renumbering of the radicals $\sqrt[p_i]{\rho_i}$ such that the first $k$ radicals satisfy

$$F(\sqrt[p_1]{\rho_1}, \sqrt[p_2]{\rho_2}, \ldots, \sqrt[p_{i-1}]{\rho_{i-1}}) \neq F(\sqrt[p_1]{\rho_1}, \sqrt[p_2]{\rho_2}, \ldots, \sqrt[p_i]{\rho_i}), \ i \leq k,$$

is easily done by taking as the first $k$ radicals those corresponding to the indices $i$ with $n_i \neq 1$ in the order as they appear in the sum defining $\gamma$. Except for their occurrence in the definition of $\gamma$ the remaining ones will not be used any further. From now on we therefore assume $n_i > 1$. Furthermore the indices $i$ will refer to this rearranged set of radicals. The first $k$ degrees $n_i$ are therefore exactly the degrees $n_i > 1$ computed above. Moreover, the order of these degrees is as in the original sequence, that is, $n_i$ is the $i$-th degree from the original sequence that is strictly larger than 1.

It follows from Theorem 3.3, p. 24, that $\sqrt[p_i]{\rho_i} \in F(i-1)$ if and only if $n_i$ divides $sd$. Equivalently, $s$ must be divisible by $\frac{n_i}{\text{gcd}(n_i, d)}$. This implies

$$E(i) = \left\{0, \frac{n_i}{\text{gcd}(n_i, d)}, 2\frac{n_i}{\text{gcd}(n_i, d)}, \ldots, (\text{gcd}(n_i, d)-1)\frac{n_i}{\text{gcd}(n_i, d)}\right\}.$$ 

Using one of the standard gcd-algorithms (for example [Sc1], but even the ordinary algorithm of Euclid suffices) shows that all sets $E(i)$ can be determined within the time bounds stated.

By Remark 6.5.1, p. 121, this step is not necessary in the complex case. It remains to compute a primitive element $\eta_i$ over $\mathbb{Q}$ for the extension $F(i)$ and the minimal polynomial $p_i$ of $\eta_i$.

Let

$$c = 2^{3\mathcal{L}}$$

we claim that

$$\eta_i = c\alpha + \sum_{j=1}^{i} \sqrt[p_j]{\rho_j}$$

is a primitive element for $F(i)$ over $\mathbb{Q}$.

By Lemma 2.5, p. 17, it suffices to show $\sigma(\eta_i) \neq \tau(\eta_i)$ for any pair $(\sigma, \tau)$ of distinct embeddings of $\mathbb{Q}(\alpha, \sqrt[p_1]{\rho_1}, \sqrt[p_2]{\rho_2}, \ldots, \sqrt[p_i]{\rho_i})$ over $\mathbb{Q}$.

First assume $\sigma(\alpha) = \tau(\alpha) = \alpha_m$. In this case

$$\sigma \left( \sum_{j=1}^{i} \sqrt[p_j]{\rho_j} \right) \neq \tau \left( \sum_{j=1}^{i} \sqrt[p_j]{\rho_j} \right)$$
has to be shown.

Consider the fields
\[ \sigma(Q(\alpha, \sqrt[n]{p_1}, \sqrt[n]{p_2}, \ldots, \sqrt[n]{p_i})) = Q(\sigma(\alpha), \sigma(\sqrt[n]{p_1}), \ldots, \sigma(\sqrt[n]{p_i})) \]
and
\[ \tau(Q(\alpha, \sqrt[n]{p_1}, \sqrt[n]{p_2}, \ldots, \sqrt[n]{p_i})) = Q(\tau(\alpha), \tau(\sqrt[n]{p_1}), \ldots, \tau(\sqrt[n]{p_i})). \]

\(\tau \sigma^{-1}\) is an isomorphism between these fields. Moreover, since \(\sigma(\alpha) = \tau(\alpha)\) it is an isomorphism that leaves the element \(\sigma(\alpha)\) fixed. Hence \(\tau \sigma^{-1}\) is an embedding of \(Q(\sigma(\alpha), \sigma(\sqrt[n]{p_1}), \ldots, \sigma(\sqrt[n]{p_i}))\) over \(Q(\sigma(\alpha))\). Since \(\sigma \neq \tau\) it is not the identity.

By Theorem 3.11, p. 31 (or Theorem 3.13, p. 32, in the complex case) and since the radicals \(\sqrt[n]{p_j}\) are linearly independent over \(Q(\alpha)\) the sum \(\sum_{j=1}^{i} \sqrt[n]{p_j}\) is a primitive element for \(Q(\alpha, \sqrt[n]{p_1}, \sqrt[n]{p_2}, \ldots, \sqrt[n]{p_i})\). By isomorphism \(\sigma\) \((\sum_{j=1}^{i} \sqrt[n]{p_j})\) is a primitive element for \(Q(\sigma(\alpha), \sigma(\sqrt[n]{p_1}), \ldots, \sigma(\sqrt[n]{p_i}))\) over \(Q(\sigma(\alpha))\). Hence the images of \(\sigma\) \((\sum_{j=1}^{i} \sqrt[n]{p_j})\) under the different embeddings of \(Q(\sigma(\alpha), \sigma(\sqrt[n]{p_1}), \ldots, \sigma(\sqrt[n]{p_i}))\) over \(Q(\sigma(\alpha))\) are pairwise distinct.

However, \(\sigma\) \((\sum_{j=1}^{i} \sqrt[n]{p_j}) = \tau\) \((\sum_{j=1}^{i} \sqrt[n]{p_j})\) implies that the images of \(\sigma\) \((\sum_{j=1}^{i} \sqrt[n]{p_j})\) under \(\tau \sigma^{-1}\) and under the identity are equal. Hence
\[ \sigma\left(\sum_{j=1}^{i} \sqrt[n]{p_j}\right) \neq \tau\left(\sum_{j=1}^{i} \sqrt[n]{p_j}\right), \]
which was to be shown.

Therefore we may assume \(\sigma(\alpha) = \alpha_k \neq \tau(\alpha) = \alpha_m\). In this case \(\sigma(\eta_i) = \tau(\eta_i)\) implies
\[ c = \frac{\sigma\left(\sum_{j=1}^{i} \sqrt[n]{p_j}\right) - \tau\left(\sum_{j=1}^{i} \sqrt[n]{p_j}\right)}{\alpha_m - \alpha_k}. \]
Since \(\sigma(\sqrt[n]{p_j})\), \(\tau(\sqrt[n]{p_j})\) are \(d_j\)-th roots of \(\sigma(p_j)\) and \(\tau(p_j)\), respectively, since \([p_j]_\infty < 2^{nL+L+2}\), and since \(k \leq \log N\) the numerator is bounded in absolute value by \(2^{nL+L+\log N+3}\) (see Lemma 5.1.11, p. 53, for a bound on \([p_j]_\infty\) and hence a bound on \(|\sqrt[n]{p_j}|_\infty\)). By the root separation bound, Lemma 5.1.3, p. 49, the denominator is bounded in absolute value from below by \(2^{-nl-n\log n}\). By choice of \(c\) and the assumption \(L > \log N\) the equality above cannot hold. This proves that \(\eta_i\) generates \(F^{(i)}\) over \(Q\).
The minimal polynomial $p_i$ is now easily computed as follows. By Lemma 5.1.7, p. 51, $|\eta_i|_\infty < 2^{3L+i+1}$. The degree of the minimal polynomial $p_i$ of $\eta_i$ over $\mathbb{Q}$ is $nN_i$, a number we already know. Hence by Lemma 7.1.1 the length $|p_i|_2$ of $p_i$ is bounded by $2^{4nN_iL}$ (which also proves the last claim of the lemma).

Using Schönhage’s algorithm (see Theorem 5.2.7, p. 62) the polynomial $p_i$ can be computed using $O(n^4N^4L)$ elementary operations on integers of size $O(n^2N^2L)$ provided an approximation to $\eta_i$ with absolute error less than

$$\epsilon = 2^{-6n^2N^2-12n^2N^2L}$$

is given\(^{20}\).

The number of operations in all applications of Schönhage’s algorithm is bounded by

$$O(n^4N^4L),$$

since

$$\sum_{i=1}^{k} N_i = N \left( 1 + \frac{1}{n_k} + \frac{1}{n_kn_{k-1}} + \ldots + \frac{1}{n_kn_{k-1} \cdots n_2} \right) \leq N \sum_{i=1}^{\log N} \frac{1}{2^i} \leq 2N.$$

Since $\eta_i = c\alpha + \sum_{j=1}^{i} \sqrt[2]{\rho_j}$ with $c = 2^{3L}$, the required approximations to $\eta_i$ for all $i$ are easily computed from approximations to $\alpha$ and $\sqrt[2]{\rho_j}$, $j = 1, 2, \ldots, i$, with absolute error less than

$$\epsilon 2^{-(3L+1)} > 2^{-6n^2N^2+13n^2N^2L}.$$

Theorem 5.4.1, p. 68 and Lemma 5.4.6, p. 71 (in the complex case Lemma 5.4.9, p. 74, has to be used) imply that the appropriate approximations to $\alpha$, $\rho_j$, $j = 1, 2, \ldots, k$, can be determined by $O(n)$ elementary operations on floating-point numbers of size $O(n^3N^2L)$ and $O(\log N \log(n^2N^2L))$ elementary operations on floating-point numbers of size $O(n^2N^2L)$. Collecting all run times now proves the lemma.

Since the time required for Schönhage’s algorithm is the dominating term we do not know how to improve asymptotically the run times of the previous

\(^{20}\)Recall $N_i \leq N$ for all $i$ and that by the time we come to this step $N$ is already known.
lemma. However, from a practical point of view the first part can be made more efficient.

The reader may wonder why we did not use one global approximation to $\alpha$, $\sqrt[\rho_j]{\rho_j}$ for the first part of the proof, rather than updating the approximations several times. The answer is that before we know $N$ which is only at the very end of this part of Step 1, the best global bound on the approximations required is polynomial in $\prod_{i=1}^{K} d_i$. But the product may be exponential in $N$. This is obviously not good enough for our purposes. In the analysis of Step 2 and Step 3, however, we will use exactly the same strategy as in the second part of the proof above, that is, we will derive a global bound on the approximations required for the algorithms. The time needed for the approximation algorithms will be determined only at the very end of the analysis.

**Remark 7.2.2** The first part of the previous lemma shows how to determine for a set of radicals over $\mathbb{Q}$ the degree of the extension generated by these radicals. But observe that in this case for the ratio tests the lattice basis reduction algorithm need not be applied. Theorem 4.1.3, p. 38, suffices. In particular, if the input set consists of $L$-bit radicals then the algorithm leading to this theorem has to be applied only to $O(LN \log N)$-bit radicals. This proves part of the run times stated in Corollary 4.2.4, p. 43.

In Step 1 also the representation of $\gamma$ with respect to $\eta_k$ has to be determined.

**Lemma 7.2.3** The representation of $\gamma$ as a linear combination over $\mathbb{Q}$ of powers of $\eta_k$ can be computed using $O(n^5N^5L + n^3N^3K)$ elementary operations on integers of size $O(n^3N^3L + nNK)$ plus the number of operations needed to compute an approximation to $\gamma$ with error $\epsilon$ less than

$$\epsilon < 2^{-(46n^3N^3L+8nNK)}.$$

Within the same run times an integer $C < 2^{10n^2N^2L+2K}$ is determined such that $C\gamma$ is an algebraic integer.

**Proof:** By assumption, an integer $C < 2^K$ exists such that $C\gamma$ is an algebraic integer. Clearly, $|C\gamma|_\infty < 2^{2K}$.

By the previous lemma the minimal polynomial $p_k$ of $\eta_k$ over $\mathbb{Q}$ has degree $nN$ and length $|p_k|_2 < 2^{4nNL}$. By Lemma 5.1.8, p. 51 the representation size of $C\gamma$ with respect to $\eta_k$ is bounded by $2^{2nN \log nN + 8n^2N^2L+2K} < 2^{10n^2N^2L+2K}$.
Since $C\gamma$ is an algebraic integer the denominator of the representation of $C\gamma$ is bounded by $|\Delta_k|$, the discriminant of $\eta_k$ (see Lemma 2.10, p. 22), and by Lemma 5.1.5, p. 50, $|\Delta_k| < 2nN \log nN + 4n^2N^2L$. Hence the representation size of $\gamma = \frac{1}{C}C\gamma$ is bounded by $22nN \log nN + 8n^2N^3L + 2K < 210nN^2L + 2K$, too.

By Theorem 5.2.6, p. 61, and by the bound on $|p_k|$ derived in the previous lemma, given an approximation to $\gamma$ with absolute error less than $2^{-(2n^2N^2 + 7nN + nN \log nN + 4n^2N^2L + 40n^3N^3L + 8nNK)} > 2^{-(46n^3N^3L + 8nNK)} > \epsilon$

its representation can be determined using $O(n^5N^5L + n^3N^3K)$ elementary operations on integers of size $O(n^3N^3L + nNK)$.

As the integer $C$ we choose the denominator of the representation determined.

By the bound on the representation size of $\gamma$ the approximation given in the proof suffices to determine in the real case whether $\gamma$ is positive. This justifies the input assumption $\sqrt[N]{\gamma} \in \mathbb{R}$ in the real case.

By assumption an integer $C$ less than $2^K$ with $C\gamma$ exists but we may not be able to determine it since we have no exact description of the ring of integers of $\mathbb{Z}[\eta_k]$, instead we only know the superset $\frac{1}{C}\mathbb{Z}[\eta_k]$.

**Remark 7.2.4** By the bound given on $K$ and by Lemma 7.1.2 if $\gamma$ is a polynomial expression in ratios of sums of radicals as described in Section 7.1, then the approximation can be determined using $O(n)$ elementary operations on floating-point numbers of size $O(n^4N^3L + m^3n^3N^3L)$ and $O(N(\log(NL) + m^2n))$ operations on floating-point number of size $O(n^3N^3L + m^3n^2N^3L)$. ♦
7.3 Description and Analysis of Step 2

In Step 2 we have to compute for each element $\gamma^{(i)}$ in the sets $D^{(i)}$ the inverse of one element denesting it. By Remark 6.5.1, p. 121, in the complex case $D^{(i)}$ contains only one element $\gamma^{(i)}$.

For an element $\gamma^{(i)}$ the procedure that determines one inverse of an element in $F_{i-1}$ denesting it has two phases. In the first phase a superset of its set of normalized admissible sequences is computed. In the second phase until we are successful or all sequences have been tested we check for any sequence whether the formula in Lemma 6.3.2, p. 109, corresponding to $\gamma^{(i)}$, the basis $\{1, \sqrt[p]{\rho}, \ldots, \sqrt[p]{\rho^{n_i-1}}\}$, and the admissible sequence of $d$-th roots of unity leads to the inverse of an element denesting $\gamma^{(i)}$ over $F_{i-1}$.

We begin with the second phase and since we want to apply the algorithm of Theorem 5.2.6, p. 61, we need to bound the coefficient size of the elements in the sets $D^{(i)}$. By Lemma 5.1.8, p. 51, the following lemma suffices to deduce such a bound.

**Lemma 7.3.1** Let $F, \sqrt[p]{\rho_1}, \sqrt[p]{\rho_2}, \ldots, \sqrt[p]{\rho_k}, \sqrt[p]{\gamma}, F_{i-1}, i=0,1,\ldots,k$, be as before. Any element $\gamma^{(i)}$ in a set $D^{(i)}$ satisfies

$$[\gamma^{(i)}]_{\infty} \leq (n_k n_{k-1} \cdots n_{i+1})^d 2^{2d(k-i)L+3K+10n^2N^2L}.$$  

Moreover, any such $\gamma^{(i)}$ is an algebraic integer.

**Proof:** We will prove the bounds by induction on $k-i$. For $k-i = 0$ (or $i = k$) the bound is true since in Step 1 an integer $C < 2^{10n^2N^2L+2K}$ is determined such that $C\gamma = \gamma^{(k)}$ is an algebraic integer (see Lemma 7.2.3) and $D^{(k)} = \{\gamma^{(k)}\}$.

Assume that each element $\gamma^{(i)} \in D^{(i)}$ satisfies

$$[\gamma^{(i)}]_{\infty} \leq (n_k n_{k-1} \cdots n_{i+1})^d 2^{2d(k-i)L+3K+10n^2N^2L}$$

and is an algebraic integer.

By the construction in the Algorithm Denesting Element any $\gamma^{(i-1)} \in D^{(i-1)}$ can be written as

$$\left(\sum_{j=0}^{n_i-1} \zeta^{(i)} \sigma_j(\sqrt[p]{\rho_i}) \sqrt[p]{\sigma_j(\gamma^{(i)})}\right)^d \sqrt[p]{\rho_i^{nd}}$$

For a precise description of the sets $D^{(i)}$ we refer to the description of the denesting algorithms at the end of Section 6.
for \( r, s \in \{0, 1, \ldots, n_i - 1\} \), \( \gamma^{(i)} \in D^{(i)} \), \( \zeta^{(j)} \) a \( d \)-th root of unity. The mappings \( \sigma_j \) are the different embeddings of \( F^{(i)} \) over \( F^{(i-1)} \). The factor \( \sqrt[sd]{\rho} \) will not occur in the complex case since in this case \( D^{(i)} \) consists of a single element (see again Remark 6.5.1).

As mentioned in the proof of Lemma 5.3.1, p. 63,
\[
\left[\sqrt[\sigma_j(\gamma^{(i)})]{\rho}\right]_\infty \leq \left[\gamma^{(i)}\right]_\infty^{\frac{1}{d}}, \quad j = 0, 1, \ldots, n_i - 1,
\]
and similarly \( \left[\sqrt[\sigma_j(\rho_i)]{\rho}\right]_\infty \leq |\rho_i|_\infty^{\frac{1}{d}} \). Since \( |\rho_i| < 2^L \) and therefore \( |\rho_i|_\infty < 2^L \) (Lemma 5.1.11, p. 53) the latter implies
\[
\left[\sqrt[\sigma_j(\rho_i)]{\rho}\right]_\infty < |\rho_i|_\infty^{\frac{1}{d}} < 2^{\frac{1}{d}} L.
\]

Combining these estimates with Lemma 5.1.7, p. 51, shows
\[
\left[\gamma^{(i-1)}\right]_\infty \leq \left( \sum_{j=0}^{n_i-1} \left[\sqrt[\rho_j]{\gamma^{(i)}}\right]_\infty^{\frac{1}{d}} \right)^d \left[\sqrt[\rho_{sd}]{\rho}_s\right]_\infty \leq
\]
\[
\leq (n_k n_{k-1} \cdots n_i)^d 2^{2d(k-1) + 3K + 10n^2N^2L},
\]
which is also correct if the factor \( \sqrt[\rho_{sd}]{\rho}_s \) is missing.

Moreover, \( \gamma^{(i)} \) is an algebraic integer hence all its conjugates and their roots are. By the input assumptions the radicals \( \sqrt[\rho_i]{\rho} \) are algebraic integers. Since the algebraic integers form a ring this shows that \( \gamma^{(i-1)} \) is also an integer, which proves the lemma. 

\[\square\]

**Corollary 7.3.2** The infinity norm \( \left[\gamma^{(i)}\right]_\infty \) of an element \( \gamma^{(i)} \) in a set \( D^{(i)} \) satisfies
\[
\left[\gamma^{(i)}\right]_\infty < 2^{10n^2N^2L + 3dL \log N + 3K}.
\]

The coefficient size \( \left[\gamma^{(i)}\right] \) with respect to the primitive element \( \eta_0^{22} \) satisfies
\[
\left[\gamma^{(i)}\right] < 2^{20n^2N^2L + 3dL \log N + 3K}.
\]

\( \eta_0 \) is defined to be \( \alpha \).
As usual the bounds on the coefficient size follow from Lemma 5.1.8, p. 51. Recall that the degree of \( F(i) \) over \( \mathbb{Q} \) is \( nN_i \leq nN \) and that the length of the minimal polynomial \( p_i \) of \( \eta_i \) is bounded by \( 2^{4n^2L} \).

For the sake of simplicity let us introduce some additional notation.

**Definition 7.3.3** Denote \( 20n^2N^2L + 3dL\log N + 3K \) by \( B \). Hence \( 2^B \) is a global bound for the infinity norm and the coefficient size with respect to \( \eta_i \) of elements in the sets \( D(i) \).

Using these notations from Definition 7.3.3 we get the following lemma which partially analyses Step 2.

**Lemma 7.3.4** Assume that \( \gamma(i) \) is an element of \( D(i) \) which is represented as a linear combination of powers of \( \eta_i \). Furthermore assume that approximations to \( \alpha, \sqrt[d]{\rho_j}, \) a primitive \( d \)-th root of unity \( \zeta_d \), and a primitive \( n_i \)-th root of unity \( \zeta_{n_i} \) with absolute error less than

\[
\epsilon < 2^{-12n^2N^2B}
\]

are given. By \( \sigma_j \) denote the different field embeddings of \( F(i) \) over \( F(i-1) \), \( j = 0, 1, \ldots, n_i - 1 \). For any sequence \( (1, \zeta^{(1)}, \ldots, \zeta^{(n_i-1)}) \) of \( d \)-th roots of unity and any \( r \in \{0, 1, \ldots, n_i - 1\} \), using \( \mathcal{O}(nN^3B) \) elementary operations on integers of size \( \mathcal{O}(nNB) \) and \( \mathcal{O}(nN^2 + N\log\log \frac{1}{\epsilon}) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon}) \) an element \( \gamma(i-1) \in F^{(i-1)} \) can be determined such that if

\[
\left( \sum_{j=0}^{n_i-1} \zeta_j^{(j)} \sigma_j(\sqrt[d]{\rho_j})^{\sqrt[d]{\sigma_j(\gamma(i))}} \right)^d \in F^{(i-1)}
\]

then it must be \( \gamma(i-1), \gamma(i-1) \) is represented as a linear combination of powers of \( \eta_{i-1} \).

Using additional \( \mathcal{O}(n^2N^2\log d) \) elementary operations on integers of size \( \mathcal{O}(dB) \) it can be decided whether the inverse of this element denests \( \sqrt[d]{\gamma(i)} \).

Finally, within the previous run times the representations with respect to \( \eta_{i-1} \) of the multiples of \( \gamma(i-1) \) with all elements in \( \{ \sqrt[d]{\rho_s} \mid s \in E(i) \} \) can be determined.

**Proof:** By the previous bounds, if the complex number corresponds to an element \( \gamma(i-1) \in F^{(i-1)} \) then the coefficient size of \( \gamma(i-1) \) is bounded by \( 2^B \). By the bounds on \( p_i \) and by Theorem 5.2.6, p. 61, given an approximation.
to this number with absolute error less than (observe that \( N \geq 2 \) and that \( B \) contains a term quadratic in \( N \))

\[
2^{-12n^2N^2+7nN+nN \log nN+4n^2N^2L+4nNB} \geq 2^{-5nNB},
\]

then the exact representation of \( \gamma^{(i-1)} \) can be reconstructed using \( \mathcal{O}(nN^3B) \) elementary operations on integers of size \( \mathcal{O}(nNB) \).

Next we show how to obtain the required approximation from the initial approximations.

By Theorem 3.9, p. 28, we can assume that \( \sigma_j \) is defined via \( \sigma_j(\sqrt[p_j]{n}) = \zeta^{\frac{1}{p_j}}n^{\frac{1}{p_j}} \). Then \( \sigma_j(\eta_i) = c\alpha + \zeta^{\frac{1}{p_j}}n^{\frac{1}{p_j}} + \cdots + \zeta^{\frac{1}{p_j}}n^{\frac{1}{p_j}}, \) with \( c = 2^L \) (for the value of \( c \) see the proof of Lemma 7.2.1).

Now the approximation to \( \alpha \) leads to an approximation to \( c\alpha \) with absolute error less than \( 2^{-12n^2N^2B+3L} \) since \( c \) is an integer known exactly. Next the approximations to \( \zeta^{\frac{1}{p_j}}n^{\frac{1}{p_j}} \), and \( \zeta_{n_i} \) lead to approximations to \( \zeta^{\frac{1}{p_j}}n^{\frac{1}{p_j}} \) with absolute error less than \( 2^{-12n^2N^2B+3L+\log N} \). The number of steps used to compute the intermediate approximations for all \( j = 0, 1, \ldots, n_i - 1 \), is bounded by \( \mathcal{O}(n_i + \log N) \in \mathcal{O}(N) \).

Since \( nN \) is an upper bound on the degree \( nN_i \) of \( F^{(i)} \) over \( Q \), \( \sigma_j(\gamma^{(i)}) = \frac{1}{j} \sum_{h=0}^{N_i-1} f_j(\eta_i)^h, f_j \in \mathbb{Z}, j = 0, \ldots, n_i, \)

As observed in the proof of Lemma 7.2.1 \( [\eta_i]_\infty < 2^{3L+i+1} \). Hence (apply Lemma 5.4.3, p. 69, part 2), with \( \alpha = \sigma_j(\eta_i), L = B, \) and \( n = nN \) the given the approximation to \( \sigma_j(\eta_i) \) with error less than \( 2^{-12n^2N^2B+3L+\log N} \) then \( \sum_{h=0}^{N_i-1} f_j(\eta_i)^h \) is approximated with absolute error less than

\[
2^{-12n^2N^2B+3L+\log N+6nNLC+2nNI+4nN+B}. 
\]

Moreover if \( \frac{1}{j} \) is approximated with absolute error less than \( \epsilon \) then \( \sigma_j(\gamma^{(i)}) = \frac{1}{j} \sum_{h=0}^{N_i-1} f_j(\eta_i)^h \) itself is approximated with absolute error less than

\[
2^{-12n^2N^2B+3L+\log N+6nNLC+2nNI+4nN+B+2} < 2^{-11n^2N^2B}. 
\]

Since \( nN \) is an upper bound on the degree \( nN_i \) of \( F^{(i)} \) over \( Q \), \( \gamma^{(i)} = \frac{1}{j} \sum_{h=0}^{N_i-1} f_i\eta_i, f_i \in \mathbb{Z}, \) and \( n_i \leq N \) numbers \( \sigma_j(\gamma^{(i)}), j = 0, 1, \ldots, n_i - 1 \), have to be approximated, \( \mathcal{O}(nN^2) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon}) \) are needed to approximate \( \sigma_j(\gamma^{(i)}), j = 0, 1, \ldots, n_i - 1. \)
shown that the approximations to \( B \) and the fact that we still have to determine whether its inverse denests if non-zero is bounded from below by \( 2^{-\sqrt{\sigma_i}} \).

Hence the imaginary part \( \sigma_i \) of \( \sqrt{\gamma_i} \) can be used.

Using the estimates of Lemma 5.4.3, p. 69, once more, it can easily be shown that the approximations to \( \sqrt[\ell]{\sigma_j(\gamma_i)} \), \( i, j = 0, 1, \ldots, n_i - 1 \), with absolute error less than \( 2^{-5n^2N^2B} \) from the approximations to \( \sigma_j(\gamma_i) \). For the estimate on the error again \( N \geq 2 \) and the fact that \( B \) contains a term quadratic in \( N \) are used\(^{23}\).

Using the estimates of Lemma 5.4.3, p. 69, once more, it can easily be shown that the approximations to \( \sqrt[\ell]{\sigma_j(\gamma_i)} \), \( i, j = 0, 1, \ldots, n_i - 1 \), with absolute error less than \( 2^{-5n^2N^2B} \) from the approximations to \( \sigma_j(\gamma_i) \). For the estimate on the error again \( N \geq 2 \) and the fact that \( B \) contains a term quadratic in \( N \) are used\(^{23}\).

Obviously, to compute this approximation from the previous ones takes only \( O(n_i + \log d) \) elementary operations on floating-point numbers of size \( O(\log \frac{1}{\epsilon}) \), which is dominated by the \( O(nN^2 + N\log \log \frac{1}{\epsilon}) \) operations on floating-point numbers of size \( O(\log \frac{1}{\epsilon}) \) we used already previously (Observe that \( B \) and hence \( \log \frac{1}{\epsilon} \) contains a term linear in \( d \)).

If the reconstruction algorithm returns the element \( \gamma_i^{(i-1)} \) in \( F^{(i-1)} \), then we still have to determine whether its inverse denests \( \sqrt[\ell]{\gamma_i} \). Equivalently, we have to check whether

\[
\sqrt[\ell]{\gamma_i^{(i-1)d-1}} \sqrt[\ell]{\gamma_i} \in F^{(i)}.
\]

But this is exactly the kind of problem we solved in Section 5 (Theorem 5.6.5, p. 92) since \( \gamma_i^{(i-1)d-1} \gamma_i \in F^{(i)} \). It follows from Corollary 7.3.2 that the approximation and reconstruction step can be done within the same time bounds as the previous step. In particular, due to our choice of \( B \) an element in \( F^{(i)} \) corresponding to \( \sqrt[\ell]{\gamma_i^{(i-1)d-1}} \gamma_i \) has coefficient size at most \( 2^B \) and our approximations to \( \alpha, \sqrt[\ell]{p_i}, \zeta, \zeta_i \) suffice to compute an approximation to \( \sqrt[\ell]{\gamma_i^{(i-1)d-1}} \gamma_i \) as required by the reconstruction step.

\(^{23}\)If the real part of \( \sigma_j(\gamma_i) \) is zero Lemma 5.4.6, p. 71, instead of Lemma 5.4.11, p. 77, can be used.
On the other hand, for the deterministic test whether the element in $F^{(i)}$ determined by the reconstruction step is identical to $\frac{d}{\sqrt{\gamma(i-1)}} \frac{\sqrt{\gamma(i)}}{d-1}$, we compute the $d$-th power of the element in $F^{(i)}$ and compare it to $\gamma(i-1) \frac{d}{\sqrt{\gamma(i)}}$. By Lemma 5.5.5, p. 83, for this step we have to spend $O(n^2 N^2 \log d)$ elementary operations on integers of size $O(dB)$ which is not covered by the previous run times.

But for this test we need to represent $\gamma(i-1) \frac{d}{\sqrt{\gamma(i)}}$ as an element in $F^{(i)}$. We therefore compute the representation of $\gamma(i-1)$ as an element of $F^{(i)}$, raise it to the $(d-1)$-st power by exact arithmetic in $F^{(i)}$, and multiply it with $\gamma(i)$. By definition of $B$, Corollary 7.3.2, and by Lemma 5.5.2, p. 82 and Lemma 5.5.5, p. 83, the time needed for this step is dominated by the previous run times.

The last statement of the lemma follows from the fact that by the results of Section 5, in particular, Theorem 5.2.6, p. 61, the representation of $\sqrt{\rho^{(i)}}$ for the smallest non-zero $j \in E^{(i)}$ can be computed within the time bounds of the lemma. Recall from the proof of Lemma 7.3.1 that $[\sqrt{\rho^{(i)}}]_{\infty} < 2B$ and hence for this reconstruction step the same analysis as for the reconstruction step for $\gamma(i-1)$ applies. Then the powers of this element and their multiples with the element determined above are computed by exact arithmetic in $Q(\alpha)$. Since in each step the coefficient size is bounded by $2B$ (see again the proof of Lemma 7.3.1 and Corollary 7.3.2, Definition 7.3.3) the claim follows from Lemma 5.5.2, p. 82.

For $i < k$ in the overall Algorithm Denesting Element the assumption that $\gamma^{(i)}$ is represented as a linear combination of powers of $\eta^i$ is justified by the lemma itself. In fact, $\gamma^{(i)}$ itself will be computed by the procedure leading to Lemma 7.3.4 and hence it will be represented as a linear combination of powers of $\eta^i$ as is required on the next level of the Algorithm Denesting Element for the computation of $\gamma^{(i-1)}$. For the case $i = k$ the assumption that $\gamma^{(k)}$ is represented as a linear combination of powers of $\eta^k$ is justified by Lemma 7.2.3. Also observe that by this lemma the coefficient size of $\gamma$ and of $C \gamma = \gamma^{(k)}$ is bounded by $2B$.

**Remark 7.3.5** The procedure of the proof of Lemma 7.3.4 shows how to determine denesting elements for $\sqrt{\gamma}$ in case $\gamma$ is an element of an arbitrary algebraic number field (recall Definition 6.2.1, p. 105). However, in this

---

24Recall again that $nN$ is an upper bound on the degree of $F^{(i)}$ over $Q$. 

147
general setting we do not know how to determine efficiently a small superset of the set of admissible sequences for $\sqrt[\gamma]{\gamma}$. Trying all possible sequences of $d$-th roots of unity leads to an algorithm whose run time is polynomial in $2^N \log d$ (\(N\) is the degree of the extension containing $\gamma$ in this case) and the remaining input parameters. But even this improves an algorithm of Landau (see [La3]) that achieves a run time that is polynomial only in $2^{Nd}$.

By Lemma 6.3.2, p. 109, applying the previous lemma to all sequences contained in a superset for the normalized admissible sequences of an element $\gamma^{(i)} \in D^{(i)}$ over the field $F^{(i-1)}$ shows how much time is needed to determine a single denesting element and the corresponding denesting set $D_{\gamma^{(i)}}$ for $\gamma^{(i)}$ over $F^{(i-1)}$. Next we show how to compute efficiently for each element in the sets $D^{(i)}$ a superset of its set of admissible sequences. In the real case the following result is used.

**Lemma 7.3.6** Let $\zeta_{n_i}$ be a primitive $n_i$-th root of unity. Assume that $\alpha$, $\sqrt[\gamma]{\rho_i}$, $\zeta_{n_i}$, are approximated with absolute error less than

$$\epsilon < 2^{-6(n^2N^4+3n^2N^4L)+5L+1}.$$ 

In the real case primitive elements and their minimal polynomials over $\mathbb{Q}$ for the fields $F^{(i)}(\zeta_{n_i})$, $i = 1, 2, \ldots, k$, can be computed using $O(n^4N^8L)$ elementary operations on integers of size $O(n^2N^4L)$ and $O(\log N)$ operations on floating-point numbers of size $O(n^2N^4L)$.

The minimal polynomials have degree less than $nN^2$ and length bounded by $2^{6nN^2L}$.

**Proof:** We analyze the time needed to compute the primitive element for one extension $F^{(i)}(\zeta_{n_i})$.

Recall that if $\phi$ denotes Euler’s $\phi$-function then the degree of $\mathbb{Q}(\zeta_{n_i})$ over $\mathbb{Q}$ is $\phi(n_i) < n_i$. Hence the degree of the extension $F^{(i)}(\zeta_{n_i})$ over $\mathbb{Q}$ is at most $nN_i\phi(n_i)$, which is bounded by $nN_i^2$ and $nN^2$. The last bound may be rather crude in general, but for $k = 1$ it is almost tight.

By the Primitive Element Theorem 2.6, p. 17, for any integer $c' \neq 0$ satisfying

$$c' \neq \frac{\eta^{(h)} - \eta^{(j)}}{\zeta_{n_i}^u - \zeta_{n_i}^v},$$

for different powers $\zeta_{n_i}^u, \zeta_{n_i}^v$ of $\zeta_{n_i}$, and conjugates $\eta^{(j)}, \eta^{(h)}$ of $\eta$, the element $\eta + c'\zeta_{n_i}$ will generate $F^{(i)}(\zeta_{n_i})$. 

148
The denominator of the ratio is bounded from below by $2^{-\log N + \log \pi}$ since this is a lower bound for the distance between two vertices in a regular $N$-gon whose circumcircle has radius 1. Using this bound, the assumption $L > \log N$, and $[\eta_i]_\infty < 2^{4L}$ it can be shown that for $c' = 2^{5L}$ the element $\eta_i + c' \zeta_{n_i}$ is a primitive element. The bound on the length of the minimal polynomial is immediate from Lemma 7.1.1, p. 130, using $[\eta_i + c' \zeta_{n_i}]_\infty < 2^{4L} + 2^{5L} < 2^{4L + 1}$.

Moreover, applying Schönhage’s reconstruction algorithm from Theorem 5.2.7, p. 62, shows that the minimal polynomial for $\eta_i + c' \zeta_{n_i}$ can be computed using $O(n^4N^8L)$ elementary operations on integers of size $O(n^4N^4L)$ provided an approximation to $\eta_i + c' \zeta_{n_i}$ with absolute error less than $2^{-6(n^2N^4 + 3n^2N^4L)}$ exists. Such an approximation can be computed easily from the given approximations by $O(\log N)$ elementary operations on floating-point numbers of size $O(\log \frac{1}{\epsilon})$.

To obtain the run time for computing all primitive elements we just have to sum up the number of elementary operations. As used already previously $\sum_{i=1}^k N_i \leq 2N$ and the number of operations is bounded by $O(n^4N^8L)$. \[\square\]

In the complex case $F^{(i)}(\zeta_{n_i}) = F^{(i)}$ since $n_i$ divides $d_i$, hence it divides $d$, which implies that $F$ contains with a primitive $d$-th root of unity also a primitive $n_i$-th root of unity.

The next lemma is the algorithmic version of Lemma 6.4.3, p. 113.

**Lemma 7.3.7** Assume that $\alpha$, $\sqrt[d_i]{\zeta_{n_i}}$, and a primitive $d$-th root of unity $\zeta_d$ are approximated with absolute error less than

$$\epsilon < 2^{-60n^3N^6L + 13n^2N^2B}.$$  

In the real case, if a primitive element for $F^{(i)}(\zeta_{n_i})$ has already been computed then for any $\sqrt[d]{\gamma^{(i)}}$, $\gamma^{(i)} \in D^{(i)}$, a superset of size $d^3$ of its set of normalized admissible sequences over $F^{(i-1)}$ can be determined using $O(d^2(n^3N^{10}L + n^3N^8B))$ elementary operations on integers of size $O(n^3N^6L + n^3N^2B)$ and $O(d^2nN^3 + N\log \log \frac{1}{\epsilon})$ elementary operations on floating-point numbers of size $O(\log \frac{1}{\epsilon})$.

It may look strange that we compute $d^3$ sequences by an algorithm whose dependence on $d$ is only quadratic. But we can represent the sequences in a more efficient form than just listing them all. For details see the following proof.

149
Proof: By Lemma 6.4.3, any normalized admissible sequence \( \left( 1, \zeta^{(1)}, \zeta^{(2)}, \ldots, \zeta^{(n_i - 1)} \right) \) is already determined by its prefix \( \left( 1, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)} \right) \). Moreover, in this situation
\[
\zeta^{(2)} \sqrt[d]{\gamma^{(i)}d^{-1}} \sqrt[2]{\sigma_2(\gamma^{(i)})} = \rho \in F(i)(\zeta_{n_i}),
\]
and
\[
\zeta^{(2(m-1))d^{-1}} \sqrt[2]{\sigma_2(m-1)(\gamma^{(i)})d^{-1}} \sqrt[2]{\sigma_2(m)(\gamma^{(i)})} = \tau^{(m-1)}(\rho) \in F(i)(\zeta_{n_i}),
\]
where \( \tau \) is the field embedding of \( F(i)(\zeta_{n_i}) \) over \( F(i-1)(\zeta_{n_i}) \) with \( \tau(\sqrt[d]{\rho_i}) = \zeta_{n_i}^{2d} \sqrt[d]{\rho_i} \). \( \tau^{(m-1)}(\rho) \) denotes \( \tau(\tau(\tau(\ldots(\tau(\rho)\ldots)) \ldots)) \), which makes sense since \( \tau \) is an automorphism of \( F(i)(\zeta_{n_i}) \) over \( F(i-1)(\zeta_{n_i}) \).

Analogously, for the odd indices we get
\[
\zeta^{(1)d^{-1}} \sqrt[d]{\sigma_1(\gamma^{(i)})d^{-1}} \sqrt[d]{\sigma_3(\gamma^{(i)})} = \rho' \in F(i)(\zeta_{n_i}),
\]
and
\[
\zeta^{(2m-1)d^{-1}} \sqrt[d]{\sigma_2(m-1)(\gamma^{(i)})d^{-1}} \sqrt[d]{\sigma_2(m+1)(\gamma^{(i)})} = \tau^{(m-1)}(\rho') \in F(i)(\zeta_{n_i}).
\]

Here as throughout the proof \( \sigma_j \) denotes the embedding of \( F(i) \) over \( F(i-1) \) determined by \( \sigma_j(\sqrt[d]{\rho_i}) = \zeta_{n_i}^{d} \sqrt[d]{\rho_i} \).

Observe that the subsequences for the even and the subsequences for the odd indices are independent of one another. Hence we will determine the set of sequences for the even and the odd indices separately. The set of admissible sequences is obtained by mixing these two sets of sequences in all possible ways.

For the even indices the following algorithm is used.

For all \( d \)-th roots of unity \( \zeta^{(2)} \) do the following:

Use the algorithm of Theorem 5.2.6, p. 61, to compute an element \( \rho \) in \( F(i)(\zeta_{n_i}) \) such that if \( \zeta^{(2)} \sqrt[d]{\gamma^{(i)}d^{-1}} \sqrt[d]{\sigma_2(\gamma^{(i)})} \in F(i)(\zeta_{n_i}) \) then it must be \( \rho \). If no such \( \rho \) exists proceed with the next \( d \)-th root of unity, since no admissible sequence with \( \zeta^{(2)} \) in the third position can exist.
Using the formula
\[ \zeta^{(2m)} = \frac{\tau^{(m-1)}(\rho)}{\sqrt[d]{\sigma(2(m-1))} \sigma^{d-1} \sqrt[2]{\sigma_{2m}(\gamma(i))}} \zeta^{(2(m-1))}, \]
determine for \( m = 1, 2, \ldots, \left\lfloor \frac{n}{2} \right\rfloor \), by a bit comparison test with all \( d \)-th roots of unity the elements \( \zeta^{(2m)} \) in the admissible sequences containing \( \zeta^{(2)} \). If for some \( m \) no \( d \)-th root of unity satisfying the equation can be found proceed with the next possible value for \( \zeta^{(2)} \).

For each tuple of \( d \)-th roots of unity \( (\zeta^{(1)}, \zeta^{(3)}) \) the roots of unity \( \zeta^{(2m-1)} \) corresponding to those admissible sequences with prefix \((1, \zeta^{(1)}, \zeta^{(2)}, \zeta^{(3)})\) for some \( d \)-th roots of unity are computed in exactly the same way except that we start with the equation
\[ \zeta^{(1)} d - 1 \zeta^{(3)} \sqrt[2]{\sigma_1(\gamma(i))} d - 1 \sqrt[2]{\sigma_3(\gamma(i))} = \rho', \]
for some \( \rho' \in F(i)(\zeta_{n_i}) \).

It remains to analyze the run time of this algorithm. First observe that although there are \( d^3 \) triples of roots of unity determining normalized admissible sequences, in our approach we have to compute only \( d + d^2 \) representations of elements in \( F(i)(\zeta_{n_i}) \).

For each pair \( \sigma_m, \sigma_{m+1} \)
\[ \left[ \zeta^{(m)} \zeta^{(m+1)} \sqrt[d]{\sigma_m(\gamma(i))} d - 1 \sqrt[d]{\sigma_{m+1}(\gamma(i))} \right] < 2^B, \]
by definition of \( B \).

Hence if \( \zeta^{(2)} \sqrt[d]{\gamma(i)} d - 1 \sqrt[2]{\sigma_2(\gamma(i))} = \rho \in F(i)(\zeta_{n_i}) \) then (see Lemma 5.1.8, p. 51, recall that \( nN^2 \) is a bound on the degree of \( F(i)(\zeta_{n_i}) \) over \( \mathbb{Q} \) and recall also the bound on the minimal polynomial of a primitive element of \( F(i)(\zeta_{n_i}) \) given in the previous lemma)
\[ [\rho] < 2^{2nN^2 \log nN^2 + 12n^2 N^4 \mathcal{L} + B} < 2^{14n^2 N^4 \mathcal{L} + B}. \]

By Theorem 5.2.6, p. 61, given an approximation to \( \zeta^{(2)} \sqrt[d]{\gamma(i)} d - 1 \sqrt[d]{\sigma_2(\gamma(i))} \) with absolute error less than (recall \( N \geq 2 \))
\[ 2^{-2(2n^2 N^4 + nN^2 \log nN^2 + 7nN^2 + 6n^2 N^4 \mathcal{L} + 56n^3 N^6 \mathcal{L} + 4nN^2 B)} \geq 2^{-60n^3 N^6 \mathcal{L} + 4nN^2 B} \]

151
the exact representation of an element \( \rho \in F^{(i)}(\zeta_n) \) such that if the complex number above is in \( F^{(i)}(\zeta_n) \) then it must be \( \rho \) can be determined by \( \mathcal{O}(n^3N^6(n^2N^4\mathcal{L} + \mathcal{B})) \) elementary operations on integers of size \( \mathcal{O}(nN^2(n^2N^4\mathcal{L} + \mathcal{B})) \).

Likewise, for roots of unity \( \zeta^{(1)}, \zeta^{(3)} \) we determine \( \rho' \in F^{(i)}(\zeta_n) \).

It can be shown exactly as in the proof of Lemma 7.3.4 that the initial approximations to \( \alpha, \sqrt{d}, \sqrt{\gamma^{(i)}}, \sqrt{\sigma_1(\gamma^{(i)})}, \sqrt{\sigma_2(\gamma^{(i)})}, \sqrt{\sigma_3(\gamma^{(i)})} \) by \( \mathcal{O}(nN + \log \log \frac{1}{\epsilon}) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon}) \). Furthermore throughout all reconstruction steps at most \( d^2 \) products \( \zeta^2 \sqrt{\gamma^{(i)}}, \sqrt{\sigma_2(\gamma^{(i)})} \) or \( \zeta^{(1)^{d-1}} \zeta^{(3)} \sqrt{\sigma_1(\gamma^{(i)})}, \sqrt{\sigma_3(\gamma^{(i)})} \) need to be computed. Hence all approximations used in these steps can be determined from the initial one by \( \mathcal{O}(d^2 + nN + \log \log \frac{1}{\epsilon}) \) elementary operations on floating-point numbers of size \( \mathcal{O}(\log \frac{1}{\epsilon}) \).

If \( \zeta^{(2(m-1))} \) is already computed, in order to determine the \( d \)-th root corresponding to \( \tau^{(m-1)}(\rho)\zeta^{(2(m-1))} \sqrt{\sigma_2(\gamma^{(i)})} d \sqrt{\sigma_3(\gamma^{(i)})} \) we just have to approximate this number with absolute error less than \( 2^{-\log d} \) and compare it with all \( d \)-th roots of unity. Since \( 2^{-\log d} \) is a separation bound for two \( d \)-th roots of unity this will pick out the correct \( d \)-th root of unity if the ratio is such a root.

We determine such an approximation by approximating the numerator and denominator separately. As mentioned above the numerator must be bounded in absolute value by \( 2^B \). By Lemma 5.1.2, p. 48, Lemma 7.1.1, p. 130, and Corollary 7.3.2 the denominator is bounded from below by \( 2^{-nNB-1} \). In fact, by Corollary 7.3.2 we know that \( [\gamma^{(i)}]_\infty < 2^B \). The degree of \( \gamma^{(i)} \) and its conjugates is at most \( nN \) and therefore by Lemma 7.1.1 its minimal polynomial has infinity norm bounded by \( 2^{nNB} \). Hence the same is true for the minimal polynomial of \( \frac{1}{\gamma^{(i)}} \). By Cauchy’s bound (Lemma 4.2.1, p. 41) \( [\gamma^{(i)}]_\infty \) lower bounded by \( 2^{-nNB-1} \). But this immediately implies that the denominator is also bounded from below by \( 2^{-nNB-1} \).

Hence approximating the numerator with absolute error less than \( 2^{-(nNB+\log d+3)} \) and the denominator with absolute error less than \( 2^{-(B+\log d+2)} \) suffices.

We know the representation of \( \rho \) as a linear combination of powers of
the primitive element $\eta_i + c'\zeta_{ni}$ and

$$
\tau^{(m-1)}(\eta_i + c'\zeta_{ni}) = \tau^{(m-1)}(\eta_i) + c'\tau^{(m-1)}(\zeta_{ni}) = \sigma_2^{(m-1)}(\eta_i) + c'\zeta_{ni}
$$

Since $\tau^{(m-1)}(\rho)$ is represented as a linear combination of $O(nN^2)$ powers of $\eta_i + c'\zeta_{ni}$ as in the proof of Lemma 7.3.4 it can be shown that the required approximation to $\tau^{(m-1)}(\rho)$ can be determined from the given approximations using $O(nN^2)$ elementary operations on floating-point numbers of size $O(\log 1/\epsilon)$ (in the proof of Lemma 7.3.4 we had to approximate $\sigma_j(\gamma^{(i)})$).

There are at most $O(d^2 N)$ different numerators so overall this step takes $O(d^2 n N^3)$ elementary operations on floating-point numbers of size $O(\log 1/\epsilon)$.

Lemma 5.4.11, p. 77, implies that the required approximations to each denominator can be computed from the initial ones by $O(n N^i + \log \log 1/\epsilon)$ elementary operations on floating-point numbers of size $O(\log 1/\epsilon)$. But observe that overall there are at most $n_i$ different denominators that have to be approximated. Hence for all ratios appearing during the algorithm this step can be done in with $O(n N^2 + N \log \log 1/\epsilon)$ elementary operations on floating-point numbers of size $O(\log 1/\epsilon)$.

Multiplying the approximations in order to approximate the ratio can clearly be done by one operation on floating-point numbers of size $O(\log 1/\epsilon)$ and throughout the algorithm at most $O(d^2 N)$ ratios need to be determined.

Finally, we needed to compute all $d - 1$ powers of the primitive $d$-th root of unity $\zeta_d$ (which is done of course only once) and compare the first $\log d$ bits of these numbers with the first $\log d$ bits of each of the ratios we determine in the algorithm. This takes $O(d^3 \log d)$ time which is dominated by time needed for $O(d^2(n^5 N^{10} L + n^3 N^6 B))$ elementary operations on integers of size $O(n^5 N^L + n N^2 B)$ since $B$ contains a term that is linear in $d$.

The lemma follows.

As mentioned at the end of Section 6.4 for $F = \mathbb{Q}$ we can compute a superset of the set of admissible sequences of size $O(n_i^3)$.

**Lemma 7.3.8** Suppose $F = \mathbb{Q}$ and that the radicals $\sqrt[n]{\rho_i}$ are real. Assume that $\sqrt[n]{\rho_i}, \zeta_{ni}$, and a primitive $d$-th root of unity $\zeta_d$ are approximated with absolute error less than

$$
\epsilon < 2^{-(125 N^{6} L + 39 d L \log N + 39 N^{2} K)}.
$$

153
If a primitive element for the field $F^{(i)}(ζ_{n_i})$ has been computed then for $\sqrt[3]{d}γ^{(i)}$, $γ^{(i)} ∈ D^{(i)}$ a superset of size $24n_i^3$ of its set of admissible sequences can be determined using $O((d + N^2)(N^{10}L + dLN^6 \log N + N^6K))$ elementary operations on integers of size $O(N^6L + dLN^2 \log N + N^2K)$ and $O(d \log dN^4)$ elementary operations on floating-point numbers of size $O(dN^4L + d^2L \log N + dK)$.

Proof: For all $d$-th roots of unity $ζ$ determine the element $ρ ∈ F^{(i)}(ζ_{n_i})$ such that if $ζ \sqrt[d]{γ^{(i)d−1}} 1/\sqrt[3]{d}σ_1(γ^{(i)}) ∈ F^{(i)}(ζ_{n_i})$ then it must be $ρ$. We do this in exactly the same way as in the previous lemma.

Determine the representation $μ_i$ in $F^{(i)}(ζ_{n_i})$ of $γ^{(i)d−1}σ_1(γ^{(i)})$ by determining separately the representation of $γ^{(i)}$ and $σ_1(γ^{(i)})$ in $F^{(i)}(ζ_{n_i})$ and by computing $γ^{(i−1)d−1}σ_1(γ^{(i)})$ with exact arithmetic in $F^{(i)}(ζ_{n_i})$.

Then compute $ρ^d$ by successive squaring and check whether $ρ^d = μ_i$. If so, approximate $ζ \sqrt[d]{γ^{(i)d−1}} 1/\sqrt[3]{d}σ_1(γ^{(i)})$ with absolute error less than $2^{−\log d}$ and check whether it is $ζ$.

By Theorem 6.4.4, p. 117, only $24n_i$ roots of unity will pass this test. Denote the set of these roots of unity by $Z_1$.

Replacing $σ_1(γ^{(i)})$ by $σ_2(γ^{(i)})$ and $σ_3(γ^{(i)})$, respectively, the sets $Z_2$ and $Z_3$ are defined and computed.

By Lemma 6.4.1, p. 112, only triples in $Z_1 × Z_2 × Z_3$ can be a prefix of an admissible sequence. So if the algorithm leading to the previous lemma is run only on these triples it will determine a superset of the set of admissible sequences of size at most $(24n_i)^3$.

The analysis of the algorithm described above is exactly as in the proof of the previous lemma except for the part of the algorithm in which we compute $γ^{(i−1)}σ_j(γ^{(i)})$, $j = 1, 2, 3$, and $ρ^d$ using exact arithmetic in $F^{(i)}(ζ_{n_i})$. This part is analyzed using Lemma 5.5.2, p. 82, and Lemma 5.5.5, p. 83. Moreover, we adjusted the run time for the case $F = Q$, i.e., $n, l$ do not appear.

Due to $F^{(i)} = F^{(i)}(ζ_{n_i})$ in the complex case we get a better result than Lemma 7.3.7. In particular, the degree of $F^{(i)}(ζ_{n_i}) = F^{(i)}$ is still bounded by $nN$ and the length of the minimal polynomial of a primitive element for this field is still $2^{4nN^2}$ (Lemma 7.2.1) rather than $2^{6nN^2}$ as in the real case (see Lemma 7.3.6). The proof is as above except that we use the complex version of Lemma 6.4.3, p. 113.
Lemma 7.3.9 Assume that $\alpha$, $\sqrt[\rho]{\gamma}$, $\zeta$ are approximated with absolute error less than

$$\epsilon \leq 2^{-51n^3 \mathcal{N}^3 + 13n^2 \mathcal{N}^2 \mathcal{B}}.$$ 

In the complex case, a superset of size $d$ for the set of admissible sequences of $\sqrt[\gamma]{\gamma^{(i)}}$ over $F^{(i-1)}$ can be determined using $O(d(n^3 \mathcal{N}^3 + n^3 \mathcal{N}^3 \mathcal{B}))$ elementary operations on integers of size $O(n^5 \mathcal{N}^5 \mathcal{L} + 13n^2 \mathcal{N}^2 \mathcal{B})$ and $O(dn^2 + N \log \log \frac{1}{\epsilon})$ elementary operations on floating-point numbers of size $O(\log \frac{1}{\epsilon})$.

7.4 Description and Analysis of Step 3

It remains to analyze Step 3 of the Algorithm Denesting Element. We will show that for each element $\gamma(0)$ in $D(0)$ the time needed for this step is fully covered by the time needed for a single application of the algorithm leading to Lemma 7.3.4, that is, neither do we use larger numbers than in this algorithm nor do we use more operations. The analysis we present is not optimal but it suffices for our purposes.

For each element $\gamma(0)$ in $D(0)$ (in the complex case there will be only one) we have to check whether $C_{\gamma(0)}$ denests $\sqrt[\gamma]{\gamma}$ and compute the corresponding element $\rho = \sqrt[\rho]{C_{\gamma(0)}} \sqrt[\rho]{\gamma}$ in $F(\alpha, \sqrt[\rho]{\rho_1}, \sqrt[\rho]{\rho_2}, \ldots, \sqrt[\rho]{\rho_k})$. Equivalently, we determine whether $C_{\gamma(0)}(d-1)$ denests $\sqrt[\gamma]{\gamma}$ and, if so, determine the representation of $\sqrt[\gamma]{C_{\gamma(0)}(d-1) \sqrt[\gamma]{\gamma}}$ as an element in $F(k)$.

This step is of course done by the algorithms leading to Theorem 5.2.6, p. 61, applied with the field $F(k) = \mathbb{Q}(\eta_k)$. We need to bound the representation size of $\sqrt[\gamma]{C_{\gamma(0)}(d-1) \sqrt[\gamma]{\gamma}}$ as an element of $F(k)$ with respect to the primitive element $\eta_k$. Since $C_{\gamma}$ is by choice of $C$ an algebraic integer and by Lemma 7.3.1 the same is true for $\gamma(0)$ and hence for $\sqrt[\gamma]{C_{\gamma(0)}(d-1) \sqrt[\gamma]{\gamma}}$. By Lemma 2.10, p. 22, this implies $\sqrt[\gamma]{\Gamma_{\gamma(0)}(d-1) \sqrt[\gamma]{\gamma}} = \frac{1}{\Delta_k} \sum_{i=0}^{n-1} e_i \eta_k^i$, $e_i \in \mathbb{Z}$, where $\Delta_k$ is the discriminant of $\eta_k$. By Corollary 7.3.2

$$|\sqrt[\gamma]{C_{\gamma(0)}(d-1) \sqrt[\gamma]{\gamma}}|_\infty < 2^{10n^2 \mathcal{N}^2 \mathcal{L} + 3d \mathcal{L} \log \mathcal{N} + 3K}$$

and by Lemma 7.3.6 the minimal polynomial of $\eta_k$ has degree at most $n\mathcal{N}^2$ and length at most $2^{6n^2 \mathcal{N} \mathcal{L}}$. Lemma 5.1.8, p. 51, shows $|e_i| < 2^B$.

Now Theorem 5.2.6, p. 61 shows that the time needed for determining the representation of $\sqrt[\gamma]{C_{\gamma(0)}(d-1) \sqrt[\gamma]{\gamma}}$ as an element of $F(k)$ is dominated by
the run times stated in Lemma 7.3.4. In fact, at the end of the algorithm leading to this lemma we used a similar procedure to determine whether the element \( \gamma^{(i-1)} \) denests \( \sqrt[\ell]{\gamma(i)} \). The approximations used in this lemma are also sufficient for the reconstruction algorithm in Step 3.

If \( C\gamma^{(0)d-1} \) denests \( \sqrt[\ell]{\gamma} \), this procedure already finds an expression for \( \sqrt[\ell]{C\gamma^{(0)d-1} \gamma} \) containing no depth 2 radicals and hence it finds a denesting for \( \sqrt[\ell]{\gamma} \). But the representation for \( \sqrt[\ell]{C\gamma^{(0)d-1} \gamma} \) is not as a linear combination of the elements \( \beta_0, \beta_1, \ldots, \beta_{N-1} \) of the standard basis for \( F(k) \).

Rather it finds a representation as a linear combination of powers of \( \eta_k = c\alpha + \sqrt[k]{\rho_1^1} + \sqrt[k]{\rho_2^2} + \ldots + \sqrt[k]{\rho_k^k} \), a primitive element of \( F(k) \) over \( \mathbb{Q} \). As it turns out for the General Denesting Algorithm this is not an appropriate form. Therefore we included in the Algorithm Denesting Element the last step in which we determine the representation of \( \sqrt[\ell]{C\gamma^{(0)d-1} \gamma} \) as a linear combination of the \( \beta_j \)'s. We describe and analyze the transformation.

To do so need a bound on the size of the coefficients \( \mu_j^{(i)} \in \mathbb{Q}(\alpha) \) in the representation \( \eta_k = \sum_{j=0}^{N-1} \mu_j^{(i)} \beta_j, i < nN \). Since \( k \leq \log N, i < nN \) (the degree of \( F(k) \) over \( \mathbb{Q} \)), if \( \eta_k \) is expanded it consists of at most \( (\log N + 1)^{nN} \) terms. All these terms are products of powers of \( c\alpha \), and of powers of \( \sqrt[k]{\rho_j^f}, f < nN \), with each exponent being smaller than \( nN \).

Using Lemma 5.1.8, p. 51, \( c < 2^{3L}, |\alpha| < 2^l \), easily shows

\[
[(c\alpha)^e] < 2^{6nLE}, \text{ for all } e < nN.
\]

We also know that each power product \( \prod \sqrt[k]{\rho_j^{f_j} \rho_j^{f_j}} = \sqrt[k]{\prod_{j=1}^{N} \rho_j^{f_j}} \), can be written as a product of an element \( \kappa \in \mathbb{Q}(\alpha) \) and some basis element \( \beta_h \) (Lemma 3.6, p. 26, in the real and Lemma 3.8, p. 27, in the complex case).

Define \( \ell_j := N/d_j \) for all \( j \). Then

\[
\prod_{j=1}^{k} \sqrt[k]{\rho_j^{f_j}} = \sqrt[k]{\prod_{j=1}^{k} \rho_j^{f_j \ell_j}}
\]

and

\[
\beta_h = \sqrt[k]{\prod_{j=1}^{k} \rho_j^{f_j}}
\]

for some integer \( e_j \) between 0 and \( n_j - 1 \) (see also the proof of Lemma 7.2.1,
p. 135, where a similar argument was used). Hence

\[ \kappa = N \sqrt[\prod_{j=1}^{\kappa} \rho_j^{\ell_j f_j}}. \]

Observe that \( k \leq \log N, \ell_j f_j < nN^2 \), and \( e_j \ell_j \leq N \). Therefore Lemma 5.5.5, p. 83, implies that \( \prod \rho_j^{e_j \ell_j} \) is an element in \( \mathbb{Z}[\alpha] \) whose coefficients are bounded in absolute value by \( 2^{L N \log N} \) and that \( \prod \rho_j^{e_j \ell_j} \) is an element in \( \mathbb{Z}[\alpha] \) whose coefficients are bounded in absolute value by \( 2^{L nN^2 \log N} \).

By Lemma 5.3.2, p. 64, if the ratio \( \kappa \) is in \( \mathbb{Q}(\alpha) \) then

\[ \left| \kappa \right| < 2^{6L nN^2 \log N}. \]

However, due to the fact that the numerator and denominator are algebraic integers and that the denominator has coefficient size \( 2^{L N \log N} \) the denominator of \( \kappa \) is even bounded by \( 2^{2L nN \log N} \). As used already several times, this implies that an integer less than \( 2^{2L nN \log N} \) exists such that the product of the inverse of the denominator of \( \kappa \) with this number is an algebraic integer. Hence the product of \( \kappa \) with this integer is an algebraic integer. Lemma 5.1.8, p. 51, implies that the denominator of \( \kappa \) is bounded in absolute value by \( \left| \Delta \right| 2^{2L nN \log N} \), where \( \Delta \) is the discriminant of \( \alpha \). Now the claim follows from the bound on \( \left| \Delta \right| \) in Lemma 5.1.5, p. 50.

Combining these estimates and observations implies that each power \( \eta_k^i \) can be represented as a sum \( \sum_{j=0}^{N-1} \mu_j^{(i)} \beta_j \) with

\[ \left| \mu_j^{(i)} \right| \leq 2^{6L nN^2 \log N + 6nN L + 2nN \log \log N} < 2^{10L nN^2 \log N}. \]

We transform the powers \( \eta_k^i, i = 0, 1, 2, \ldots, nN - 1 \), successively. So assume \( \eta_k^{i-1} \) has already been transformed into

\[ \eta_k^{i-1} = \sum_{j=0}^{N-1} \mu_j^{(i-1)} \beta_j, \mu_j^{(i-1)} \in \mathbb{Q}(\alpha), \left| \mu_j^{(i-1)} \right| \leq 2^{10L nN^2 \log N}. \]

\(^{25}\)Observe that a similar argument was used in the analysis of Step 1. But in this step the standard basis for \( F(k) \) may not have been computed in the form we need it now.
Hence
\[ \eta^i_k = c \alpha \sum_{j=0}^{N-1} \mu^{(i-1)}_j \beta_j + d \sqrt{\rho_1} \sum_{j=0}^{N-1} \mu^{(i-1)}_j \beta_j + \ldots + d \sqrt{\rho_k} \sum_{j=0}^{N-1} \mu^{(i-1)}_j \beta_j. \]

Each product \( d \sqrt{\rho_m} \beta_j \) must be a multiple of some element \( \kappa_{m,j} \) in \( \mathbb{Q}(\alpha) \) and a basis element \( \beta^{(j)}_m \). For each product that is not already a basis element we compute \( \kappa_{m,j} \) and \( \beta^{(j)}_m \). Then we multiply \( \kappa_{m,j} \) and \( \mu^{(i-1)}_j \), or \( c \alpha \) and \( \mu^{(i-1)}_j \), and collect the coefficients of each basis element \( \beta_j \).

The coefficients (which are elements in \( \mathbb{Q}(\alpha) \)) obtained in this way need not be in a reduced form, that is, the gcd of the denominator and the rational integers appearing in the linear combination of powers of \( \alpha \) that forms the numerator need not be 1. Accordingly, these numbers need not satisfy the bound we derived above for the coefficient size of the \( \mu^{(i)}_j \)'s. To obtain a representation satisfying this bound in the final step for each coefficient we compute the gcd of its denominator and all rational integers appearing in its numerator and divide these numbers by the gcd. Since the \( \beta_j \)'s form a basis and, accordingly, the representation
\[ \eta^i_k = \sum_{j=0}^{N-1} \mu^{(i)}_j \beta_j \]
is unique, this finally must lead to a representation with coefficient size as required.

To determine for a product \( d \sqrt{\rho_m} \beta_j \) the elements \( \kappa_{m,j} \) and \( \beta^{(j)}_m \) mentioned above we simply check for each basis element \( \beta_h \) whether the ratio \( d \sqrt{\rho_m} \beta_j / \beta_h \) is an element of \( \mathbb{Q}(\alpha) \). We do so by transforming both numerator and denominator into an \( N \)-th root as described above and apply the deterministic algorithm leading to Theorem 5.6.2, p. 90, (or Theorem 5.6.3, p. 90, in the complex case) to this ratio of radicals.

The transformation step has been analyzed already for Step 1 (see the proof of Lemma 7.2.1, p. 135). It takes \( \mathcal{O}(n^2 N \log N) \) elementary operations on integers of size \( \mathcal{O}(LN \log N) \) to transform all basis elements \( \beta_j \) and all products \( d \sqrt{\rho_m} \beta_j \) into this form.

For each ratio the algorithm leading to Theorem 5.6.2, p. 90, require \( \mathcal{O}(Ln^5 N^2 \log N) \) operations on integers of size \( \mathcal{O}(Ln^3 N^2 \log N) \), \( \mathcal{O}(n^2 \log N) \) operations on integers of size \( \mathcal{O}(Ln^2 N^3 \log N) \) (see also Table 1, page 88), and (see Theorem 5.2.6, p. 61) an approximation to the ratio with absolute error less than
\[ 2^{-25} \sqrt{Ln^3 N^2 \log N}. \]
As is easily seen (recall Lemma 5.4.3, p. 69) the approximations to $\alpha$, $\sqrt[n]{\rho_i}, i = 1, 2, \ldots, k$, used in Lemma 7.3.4, p. 144, are sufficient to determine this approximation using $O(N)$ elementary operations on floating-point numbers of size $O((\log N)^2)$. In fact, for each ratio we only need to compute $O(N)$ products of the radicals $\sqrt[n]{\rho_i}$ and perform one inversion to approximate $\frac{1}{\beta_h}$.

Moreover, observe that if we keep a list of the products $a\sqrt[n]{\rho_m}\beta_j$ that have already been computed, and every time we have to compute the basis representation of a product $a\sqrt[n]{\rho_m}\beta_j$, we check in this list whether it has already been transformed previously, we have to apply this test to at most $N \log N$ products.

So in order to compute all powers of $\eta_k$, we have to check at most $N^2 \log N$ ratios. Since the list has length at most $N \log N$, has to be updated at most $N \log N$ times, and overall $nN^2 \log N$ queries are performed the operations above clearly dominate these list operations and the time required for all ratio tests is dominated by the run time stated in Lemma 7.3.4, p. 144.

Throughout the computation of all powers of $\eta_k$ multiplying $\alpha$ or the elements $\kappa_{m,j}$ with the elements $\mu_{j}^{(i-1)}$, requires $O(n^3 N^2 \log N)$ operations on integers of size $O((\log N)^2 N)$. In fact, for each power of $\eta_k$ we have to perform $O(N \log N)$ arithmetic operations in $\mathbb{Q}(\alpha)$. Since there are $nN$ products the upper bound on the number of operations follows from Lemma 5.5.2, p. 82. The bound on the size of the integers involved in these operations follows from the bound on $[\mu_{j}^{(i)}]$.

Throughout the algorithm collecting the coefficients and computing the gcd’s uses $O(n^2 N \log N + n^2 N^2 \log((\log N)^2 N))$ elementary operations on integers of size $O((\log N)^2 N)$. In fact, for each power $\eta_k$ we first add for each basis element the coefficients. This uses $O(nN \log N)$ operations on integers of size $O((\log N)^2 N)$. To prove these bounds observe that for a fixed radical $a\sqrt[n]{\rho_m}$ any two products $a\sqrt[n]{\rho_m}\beta_i$ and $a\sqrt[n]{\rho_m}\beta_j$, $i \neq j$, must be multiples of different basis elements. Otherwise $\beta_i$ and $\beta_j$ were linearly dependent. Hence for each basis element $\beta_i$ we have to add at most $\log N$ elements in $\mathbb{Q}(\alpha)$. This shows the upper bound on the number of operations. For the bound on the integers involved recall the $O((\log N)^2 N)$ bound for the denominators of any ratio. Hence taking the product of these denominators yields an integer in $O((\log N)^2 N)$. Now the upper bound on the integers is immediate.

Then we have to compute $N$ times the gcd of $n + 1 \mathbb{O}(nN^2 \log N)$-bit integers. The run time above follows now from the fast gcd-algorithm.
Although even with the usual gcd-algorithm the run time is dominated by the run times of Lemma 7.3.4, p. 144.

Finally, given the representations of all powers of \( \eta_k \) we compute the representation of \( \sqrt[d]{C\gamma(0)} \). This step requires only arithmetic in \( \mathbb{Z} \) and can be done by \( O(n^2 N^2) \) arithmetic operations on integers of size \( O(L_n^2 N^2 \log N + dL \log N + K) \), which follows from the bounds above, and the bound on the representation size of \( \sqrt[d]{C\gamma} = \sqrt[d]{\gamma(k)} \) as a linear combination of powers of \( \eta_k \) (see Corollary 7.3.2, p. 143). So this step is also covered by the previous run times.

**Lemma 7.4.1** It can be decided for each \( C\gamma(0)^{d-1}, \gamma(0) \in D(0) \) whether it denests \( \sqrt[d]{\gamma} \) using the same approximations to \( \alpha, \sqrt[d]{\rho_i}, i = 1, 2, \ldots, k \), as in Lemma 7.3.4. Also the run times for the test is dominated by the run times stated in Lemma 7.3.4. Moreover within these time bounds the representation of \( \sqrt[d]{C\gamma(0)^{d-1} \gamma} \) as \( \sum_{j=0}^{N-1} \kappa_j \beta_j \), can be determined. Here \( \{\beta_0, \beta_1, \ldots, \beta_{N-1}\} \) is the standard basis of \( F(\sqrt[d]{\rho_1}, \ldots, \sqrt[d]{\rho_k}) \) and \( \kappa_j \in \mathbb{Q}(\alpha) \) satisfies

\[
[k_j] \in O(B) = O(Ln^2 N^2 \log N + dL \log N + K).
\]

Let us remark that combining the results of the previous subsections yields an efficient algorithm to transform any radical expression into a sum of radicals.

In fact, in the first step we determine the standard basis and a primitive element for the radical extension generated by the radicals appearing in the expression. This is done as in **Step 1**. Then we determine the representation of the expression with respect to powers of this primitive element. This is done as in **Step 2**. Then we transform this representation into a linear combination of the elements of the standard basis. This is done as in **Step 3**.
7.5 The Final Results

If we collect in the real case the run times for all applications of the single steps of the Algorithm Denesting Element we arrive at the following table. An explanation of the various entries will be given below.

**Table 2: Run Times for the Algorithm Denesting Element**
(The real case)
<table>
<thead>
<tr>
<th>procedure</th>
<th>number of operations</th>
<th>bit size of numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fields $F^{(i)}$, prim. elements in Step 1</td>
<td>$O(nN^2 \log(N\mathcal{L}))$</td>
<td>$O(n^3 N^2 \mathcal{L})$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(n^4 N^4 \mathcal{L})$</td>
<td>$O(Ln^2 N^2 \log N)$</td>
</tr>
<tr>
<td>Representation of $\gamma$ in Step 1</td>
<td>polynomial</td>
<td>polynomial (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(n^5 N^5 \mathcal{L} + n^3 N^3 \mathcal{K})$</td>
<td>$O(n^3 N^3 \mathcal{L} + nN\mathcal{K})$</td>
</tr>
<tr>
<td>Initial approximations in Step 2</td>
<td>$O(n)$</td>
<td>$O(n^4 N^6 \mathcal{L} + n^4 N^4 \mathcal{B})$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(\log N \log(nNB))$</td>
<td>$O(n^3 N^6 \mathcal{L} + n^2 N^2 \mathcal{B})$ (fp)</td>
</tr>
<tr>
<td>Prim. element $F^{(i)}(\zeta_{ni})$</td>
<td>$O(\log N)$</td>
<td>$O(n^2 N^4 \mathcal{L})$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(n^4 N^8 \mathcal{L})$</td>
<td>$O(n^2 N^4 \mathcal{L})$</td>
</tr>
<tr>
<td>Computing admissible sequences</td>
<td>$O(d^2 nN^4 + N^2 \log(nNB))$</td>
<td>$O(n^3 N^6 \mathcal{L} + n^2 N^2 \mathcal{B})$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(d^2 n^5 N^{11} \mathcal{L} + d^2 n^3 N^7 \mathcal{B})$</td>
<td>$O(n^3 N^6 \mathcal{L} + nN\mathcal{B})$</td>
</tr>
<tr>
<td>Computing denesting elements</td>
<td>$O(d^3 nN^3 + d^3 N^2 \log(nNB))$</td>
<td>$O(n^2 N^2 \mathcal{B})$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(d^3 n^3 N^4 \mathcal{B})$</td>
<td>$O(nNB)$</td>
</tr>
<tr>
<td></td>
<td>$O(d^3 n^2 N^3 \log d)$</td>
<td>$O(d\mathcal{B})$</td>
</tr>
</tbody>
</table>

By numbers both integers and floating-point numbers are meant. The symbol (fp) indicates where floating-point numbers are used. The first entry for the “Representation of $\gamma$” refers to the assumption that approximations
to $\gamma$ can be determined efficiently (see assumption 5) of Section 7.1).

To deduce these run times use Lemma 7.2.1, p. 135, Lemma 7.2.3, p. 140, Lemma 7.3.4, p. 144, Lemma 7.3.6, p. 148, and Lemma 7.3.7, p. 149, finally recall that overall the set of admissible sequences has to be determined for at most $|\bigcup_{i=1}^k D^{(i)}| \leq N$ elements (see page 120). Accordingly, the algorithm leading to Lemma 7.3.4 has to be applied at most $2d^3N$ times.

Step 3 has to be applied at most $N$ times since $|D^{(0)}| \leq N$.

As follows from Section 5.4 the run times stated above for the initial approximation step are valid to compute approximations sufficient for all applications of the algorithms leading to Lemma 7.3.4, p. 144, to Lemma 7.3.6, p. 148, and to Lemma 7.3.7, p. 149. This proves the run times in the table.

The following theorem states explicit upper bounds on the number of operations used and the bit size of the numbers on which these operations have to be performed. The bounds stated are obtained from Table 2. We simply took the worst dependence on each parameter for the number of operations as well as for the size of the integers or floating-point numbers.

Recall that

$$B = 20n^2N^2L + 3dL \log N + 3K.$$  

**Theorem 7.5.1** Let $F = \mathbb{Q}(\alpha)$ be a real algebraic number field and let $\gamma$ be a rational expression over $F$ in linearly independent radicals $d_i\sqrt[\rho_i], \rho_i \in F$, $i = 1, 2, \ldots, K$. Here $\rho_i$ an algebraic integer and $d_i\sqrt[\rho_i]$ is of degree $d_i$ over $F$. Assume $|\gamma|_\infty < 2^K$ and assume that a guarantee is given that a positive integer $C$ less than $2^K$ exists such that $C\gamma$ is an algebraic integer. Finally assume that for any positive $\epsilon$ the number $\gamma$ can be approximated with error less than $\epsilon > 0$ in time polynomial in $\log \frac{1}{\epsilon}$ and $K$. Then the algorithm Denesting Element determines whether there exists an element $\gamma^{(0)} \in \mathbb{Q}(\alpha)$ with

$$\sqrt[\rho_i]{\gamma^{(0)}} \in F(\sqrt[\rho_1]{\alpha}, \sqrt[\rho_2]{\alpha}, \ldots, \sqrt[\rho_K]{\alpha})$$

in time polynomial in $K$, the degree $n$ of the minimal polynomial $p$ of $\alpha$, in $l = \lceil \log |p|_2 \rceil$, in $L = \max\{\log|\rho_i|\}$, in the degree $N$ of the extension generated by $\sqrt[\rho_i]{\alpha}$, $i = 1, \ldots, K$, and in $d$.

In particular, the algorithm uses at most $O(d^4n^5N^{11}(L + K))$ elementary operations on integers or floating-point numbers of size $O(d^3n^5N^6(L + K))$ plus the number of operations used to compute an approximation to $\gamma$ with absolute error less than $2^{-(46n^3N^3L + 8nNK)}$. Here $L = \lceil n \log n + nl + nL \rceil$.

Using at most $N$ times this number of operations on integers or floating-point numbers of asymptotically the same size the Algorithm Real Denesting
determines a denesting of $\sqrt[\gamma]{\gamma}$ using real radicals over $\mathbb{Q}(\alpha)$, if any such denesting exists.

**Proof:** The first claim follows from Table 2.

To prove the statement on the run time of the *Algorithm Real Denesting* observe that Step 1 of the *Algorithm Denesting Element* is applied only once. Likewise the primitive elements for the fields $F(i)(\zeta_n)$ are only computed once. Step 2 and Step 3 of the *Algorithm Denesting Element* are applied to all products $\beta_j\gamma$, where $\beta_j$ is a basis element of the standard basis of $F(\sqrt[n_1]{\rho_1}, \sqrt[n_2]{\rho_2}, \ldots, \sqrt[n_K]{\rho_K})$ over $F$. Hence Step 2 and Step 3 of the *Algorithm Denesting Element* have to be applied at most $N$ times.

Observe that $\beta_j$ is an integer, hence for any rational integer $C$ such that $C\gamma$ is an integer $C\beta_j\gamma$ is an integer. Next observe that $|\beta_j\gamma|_{\infty} < 2^{L+L\log N}$ since any product $\sqrt[n_i]{\rho_{i_i}}$, $e_i < d_i$, has infinity norm less than $2^{n_i+L+2}$ (Lemma 5.1.11, p. 53) and each basis element is the product of at most $\log N$ of these powers. In fact, changing $B$ from $20n^2N^2L + 3d\log N + 3K$ to $21n^2N^2L + 3d\log N + 3K$ will already do. This will increase the corresponding bounds in Lemma 7.3.1 and in Corollary 7.3.2 only by a constant factor. Hence the value $B$ defined in Definition 7.3.3 has to be increased only by a constant factor. It follows that for each product $\beta_j\gamma$ the run times stated in Table 2 asymptotically remain unchanged.

Finally observe that in order to compute the representation of $\beta_j\gamma$ as a linear combination of powers of $\eta_k$ we may compute the representation of $\beta_j$ and $\gamma$ separately and then multiply. Computing the representation of $\gamma$ is contained in Table 2. Computing the representation of $\beta_j$ is easily dominated by determining the representation of the elements $\gamma^{(i)}$ in the intermediate denesting sets $D^{(i)}$. The second claim follows.

Recall from Section 6.2 that the denesting computed by the *Algorithm Denesting Element* is not restricted to the specific $d$-th root we denote by $\sqrt[\gamma]{\gamma}$. If $d$ is even, it applies to both $d$-th roots if it applies to any. Similar remarks apply below. Also recall that the assumptions that each $\rho_i$ is an integer and that $\sqrt[\rho_i]{\rho_i}$ is of degree $d_i$ over $F$ are no restriction. They are easily guaranteed by the algorithms of Section 5.

The claim that the algorithm runs in polynomial time without any reference to a specific model of computation is justified by the analysis in term of elementary operations. In any reasonable model of computation that allows bit operations the algorithm will use only polynomially many bit operations since it uses only polynomially many elementary operations and the
bit size on which these operations have to be performed is also polynomially bounded. The statements of the other theorems and corollaries in this section are justified similarly.

Let us explicitly state the upper bound for nested radicals of the form described in the preliminaries.

**Corollary 7.5.2** Let $\mathbb{Q}(\alpha)$ be as above. Assume $S_j$, $j = 1, 2, \ldots, m$, have the form

$$S_j = \sum_{i=1}^{m} \frac{P_{i,j}}{Q_{i,j}},$$

where $P_{i,j}, Q_{i,j}$, $i = 1, \ldots, m$, are sums of radicals of the linearly independent radicals $\sqrt[k]{\rho_i}$, $i = 1, 2, \ldots, K$, with coefficients $\kappa$ in $\mathbb{Q}(\alpha)$ satisfying $|\kappa| < 2^L$. Assume that $\rho_i$ is an integer for all $i$. Let $P(X_1, X_2, \ldots, X_m)$ be a polynomial in $m$ variables with coefficients in $\mathbb{Q}(\alpha)$. Assume that $P$ has at most $m$ terms, that the total degree is also bounded by $m$, and that its coefficients $\mu \in \mathbb{Q}(\alpha)$ satisfy $|\mu| < 2L$.

If the Algorithm Denesting Element is applied to $\gamma = P(S_1, S_2, \ldots, S_m)$ then its run time is bounded by a polynomial in $m, n, l, N, L$, and $d$. Here $N$ is the degree of the extension $\mathbb{Q}(\alpha, \sqrt[d]{\rho_1}, \sqrt[d]{\rho_2}, \ldots, \sqrt[d]{\rho_K})$ over $\mathbb{Q}(\alpha)$.

In particular, the algorithm uses at most $O(d^4 m^3 n^6 N^{13} L)$ elementary operations on integers or floating-point numbers of size $O(d^2 m^3 n^6 N^8 L)$, $L = \lceil n \log n + nl + nL \rceil$.

If the Algorithm Real Denesting is applied to $\sqrt[n]{\gamma}$ it uses at most $N$ times this number of operations on numbers of the same bit size.

**Proof:** By the results in Section 7.1 in this case $K \in O(m^3 n^2 L)$. Furthermore recall from Remark 7.2.4 that the required approximation to $\gamma$ can be determined by $O(n)$ elementary operations on floating-point numbers of size $O(n^4 N L + m^3 n^3 N^3 L)$ and $O(N \log(N L) + m^2 n)$ operations on floating-point number of size $O(n^3 N^3 L + m^3 n^2 N^3 L)$.

Setting $n = 1$ and $l = 0$ one easily extracts from these results the result for $F = \mathbb{Q}$ and real radicals $\sqrt[n]{q_1}, \ldots, \sqrt[n]{q_K}$ such that the sum of the bit size of the numerator and denominator in each $q_i$ is less than $L$. However, if Lemma 7.3.8, p. 153, instead of Lemma 7.3.7, p. 149, is used the algorithm may be more efficient. We state the result only for radical expressions as in the previous corollary.
Corollary 7.5.3 If \( \mathbb{Q}(\alpha) = \mathbb{Q} \) and \( \sqrt[\bar{\gamma}]{\bar{\gamma}} \) is as in Corollary 7.5.2 then the Algorithm Denesting Element uses \( \mathcal{O}(d^2m^3\bar{n}^3N^{13}L) \) on integers and floating-point numbers of size \( \mathcal{O}(d^2m^3N^{8}L) \).

Using at most \( N \) times this number of operations on numbers of the same size the Algorithm Real Denesting determines a denesting using real radicals if any such denesting exists.

Here \( \bar{n} \) is the maximum of all field degrees \( [\mathbb{Q}(\sqrt[\bar{\gamma}]{\bar{\gamma}_1}, \ldots, \sqrt[\bar{\gamma}]{\bar{\gamma}_i}) : \mathbb{Q}(\sqrt[\bar{\gamma}_1], \ldots, \sqrt[\bar{\gamma}_{i-1}])] \).

Observe that depending on \( k \) the maximum degree \( \bar{n} \) may be anything between 2 and \( N \). If \( k = 1 \), \( \bar{n} = N \), and \( N \) is large compared to \( d \) it may be preferable not to use the algorithm of Lemma 7.3.8. But if \( k = \log N \) and \( d \) is large applying Lemma 7.3.8 improves the run time considerably. In particular, it reduces the number of applications of the lattice basis reduction algorithm in the procedure where denesting elements are determined.

In the complex case we get a similar table than the one above. Due to Lemma 7.3.9, p. 155, the fact that the number of admissible sequences that are computed is bounded by \( d \), and the fact that on each level from \( k - 1 \) to 0 instead of computing the inverses of a denesting set we only need to determine the inverse of a single denesting element the run times are much better. Recall that \( k \) and hence the number of levels is bounded by \( \log N \) then apply Lemma 7.2.1, p. 135, Lemma 7.2.3, p. 140, Lemma 7.3.4, p. 144, and Lemma 7.3.7, p. 149.

**Table 3: Run Times for the Algorithm Denesting Element**
(The complex case)
<table>
<thead>
<tr>
<th>procedure</th>
<th>number of operations</th>
<th>bit size of numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field $F^{(i)}$, prim. elements in Step 1</td>
<td>$O(nN^2 \log(nL))$</td>
<td>$O(n^3N^2L)$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(n^4N^4L)$</td>
<td>$O(Ln^2N^2 \log N)$</td>
</tr>
<tr>
<td>Representation of $\gamma$ in Step 1</td>
<td>polynomial</td>
<td>polynomial (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(n^5N^5L + n^3N^3K)$</td>
<td>$O(n^3N^3L + nNK)$</td>
</tr>
<tr>
<td>Initial approximations in Step 2</td>
<td>$O(n)$</td>
<td>$O(n^3N^3L + n^2N^2B)$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(\log N \log(nNB))$</td>
<td>$O(n^3N^3L + n^2N^2B)$ (fp)</td>
</tr>
<tr>
<td>Computing admissible sequences</td>
<td>$O(dn^2N^2 \log N + N \log N \log(nNB))$</td>
<td>$O(n^3N^3L + n^2N^2B)$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(dn^3N^5L \log N + +dn^3N^3B \log N)$</td>
<td>$O(n^3N^3L + nNB)$</td>
</tr>
<tr>
<td>Computing denesting elements</td>
<td>$O(dn^2N^2 \log N + +dN \log N \log(nNB))$</td>
<td>$O(n^2N^2B)$ (fp)</td>
</tr>
<tr>
<td></td>
<td>$O(dn^3N^3B \log N)$</td>
<td>$O(nNB)$</td>
</tr>
<tr>
<td></td>
<td>$O(dn^2N^2 \log d \log N)$</td>
<td>$O(dB)$</td>
</tr>
<tr>
<td>Step 3</td>
<td>covered by Step 2</td>
<td></td>
</tr>
</tbody>
</table>

Recall that in the complex case the Algorithm Denesting Element already determines whether a denesting using radicals of the form $\sqrt[d]{\rho}$, $t|d$, exists.

**Theorem 7.5.4** Assume the algebraic number field $F = \mathbb{Q}(\alpha)$ contains a primitive $d$-th root of unity. Assume $\gamma$ is a rational expression over $F$ in linearly independent radicals $\sqrt[d]{\rho_i}$, $[\rho_i] < 2^L$, $i = 1, 2, \ldots, K$. Here $\rho_i$ an algebraic integer which is of degree $d_i$, $d_i|d$ over $F$. Assume $[\gamma]_\infty < 2^K$ and assume that a guarantee is given that a positive integer $C$ less than $2^K$ exists such that $C\gamma$ is an algebraic integer. Finally assume that for any positive
\( \epsilon \) the complex number \( \gamma \) can be approximated with error less than \( \epsilon > 0 \) in time polynomial in \( \log \frac{1}{\epsilon} \) and \( K \).

Then the Algorithm Denesting Element determines in time polynomial in \( n, l, N, L, d, K \) a denesting of \( \sqrt[\gamma]{\gamma} \) using radicals of the form \( \sqrt[\gamma]{\rho} \), \( l(d, \rho) \in \mathbb{Q}(\alpha) \), if any such denesting exists. The parameters \( n, l, N \) are defined as in Theorem 7.5.1.

In particular, the algorithm uses at most \( O(d^2n^5N^5(K + L)\log N) \) elementary operations on integers or floating-point numbers of size \( O(d^2n^5N^3(L + K)) \) plus the number of operations used to compute an approximation to \( \gamma \) with absolute error less than \( 2^{-(46n^3N^3L + 8nNK)} \).

Again the bounds given above are very crude and more precise run times can be read of from Table 3.

Thanks to the fact that \( \mathbb{Q}(\alpha) \) contains a primitive \( d \)-th root of unity, if the denesting applies to that \( d \)-th root of \( \gamma \) we usually denote by \( \sqrt[\gamma]{\gamma} \) then it applies to all \( d \)-th roots of \( \gamma \). As before the assumptions on the radicals are no restriction.

As we mentioned before, the last theorem can also be applied to nested radicals if \( d_i \) does not divide \( d \). In that case, the field \( \mathbb{Q}(\alpha) \) should contain primitive \( d \)-th and \( d_i \)-th roots of unity. In this situation the Algorithm Denesting Element finds a denesting using only radicals of degree dividing \( \text{lcm}(d, d_1, d_2, \ldots, d_K) \).

Corollary 7.5.2 generalizes accordingly to complex radicals.

**Corollary 7.5.5** If \( \gamma \) is of the form described in Corollary 7.5.2 then the Algorithm Denesting Element in the complex case uses \( O(d^2m^3n^6N^7L\log N) \) elementary operations on integers or floating-point numbers of size \( O(d^2m^3n^6N^6L) \).

The analysis of the General Denesting Algorithm requires still some work.

**Theorem 7.5.6** Suppose \( S = \sum_{i=1}^{K} \kappa_i \sqrt[\gamma_i]{\gamma_i} \) is a real depth two radical expression over a real algebraic number field \( \mathbb{Q}(\alpha) \), where \( \kappa_i, \gamma_i \) are radical expressions satisfying the assumptions of Theorem 7.5.1. In particular their infinity norm is bounded by \( 2^K \) and for each expression \( \gamma_i, \kappa_i \) an integer \( C_{\gamma_i} \), \( C_{\kappa_i} < 2^K \) exists such that \( C_{\kappa_i} \kappa_i \) or \( C_{\gamma_i} \gamma_i \) is an algebraic integer.

Denote by \( n \) the degree of the minimal polynomial \( p \) of \( \alpha \), \( l = \lceil \log |p|_2 \rceil \), \( 2^L \) is the maximum coefficient size of an element in \( \mathbb{Q}(\alpha) \) appearing in the definition of \( \kappa_i, \gamma_i \), and \( d \) is the maximum over all products \( d_id_j \) for indices
i, j between 1 and k. Finally assume that for any fixed pair of indices i, j the radicals appearing in \( \kappa_i, \kappa_j, \gamma_i, \) and \( \gamma_j \) generate an extension of \( \mathbb{Q}(\alpha) \) of degree at most \( N \).

When applied to \( S \) the General Denesting Algorithm determines in time polynomial in \( k, n, l, L, N, d \), and the upper bound \( K \) whether \( S \) can be written as a sum of real depth one radicals over \( \mathbb{Q}(\alpha) \). In particular, the General Denesting Algorithm uses at most \( O(k^2d^4n^6L^{12}(L + dK)) \) elementary operations on integers or floating-point numbers of size \( O(kd^2n^5L + 12dnK) \) plus the number of operations needed to approximate each \( \gamma_i, \kappa_i \) with absolute error less than \( 2^{-(46n^3L+12dnK)} \).

Within the same time bounds the algorithm decides whether \( S = 0 \).

**Proof:** For a pair of indices \( i, j \) denote the radical extension generated by the radicals in \( \gamma_i, \kappa_i, \gamma_j, \kappa_j \) by \( E_{ij} \).

Recall that we first partition \( R = \{ \sqrt[2]{\gamma_1}, \ldots, \sqrt[k]{\gamma_k} \} \) into subsets \( R_1, \ldots, R_h \) such that two nested radicals are in the same subset if and only if their ratio denests using real radicals. We assume w.l.o.g. that \( \sqrt[2]{\gamma_t} \in R_t, t = 1, \ldots, h \). Then we transform \( S \) into

\[
S = \sum_{t=1}^{h} \frac{\sqrt[2]{\gamma_t}}{\gamma_t} \sum_{\{i \in \mathbb{N} | \sqrt[k]{\gamma_i} \in R_t\}} \kappa_i \gamma_{it},
\]

\( \gamma_{it} \) denotes the denesting of \( \sqrt[k]{\gamma_i} \).

Since \( \sqrt[k]{\gamma_i} \) is an element of a radical extension of degree at most \( N \) for each of these \( \frac{k(k-1)}{2} \) expressions we can easily determine its denesting using the Algorithm Real Denesting. In particular, observe that any product has infinity norm less than \( 2^{dk} \). Hence replacing in Lemma 7.3.1, p. 142, and in the definition of \( B \) (see Definition 7.3.3, p. 144) \( K \) by \( dK \) Table 2 or Theorem 7.5.1 easily yields the run times. Moreover observe that in Lemma 7.2.3, p. 140, \( K \) also has to be replaced by \( dK \). Finally, observe that by Lemma 5.4.3, p. 69, approximations to \( \gamma_i, \gamma_t \) with absolute error less than \( 2^{-(46n^3L+12dnK)} \) suffices to determine the approximation to \( \sqrt[k]{\gamma_i} \) as required by Lemma 7.2.3.

Remark that in this step a primitive element \( \eta_{it} \) for the extension generated by the radicals in \( \gamma_i, \gamma_t \) has been computed. We may assume that this extension is already \( E_{it} \).

The next step that checks which of the radicals \( \sqrt[k]{\gamma_i}, i = 1, 2, \ldots, h \), denests is also easily done within the time bounds stated. Recall that there
is at most one radical that denests and we assume that $\sqrt[\gamma_1]{\gamma}$ is this one, if any.

In the final step we have to check whether the sums

$$S_t = \sum_{\{i \in \mathbb{N} \mid \sqrt[\gamma_i]{\gamma_i} \in R_t\}} k_i \gamma_{it},$$

for $t \geq 2$, if $\sqrt[\gamma_1]{\gamma}$, denests, and for $t \geq 1$, otherwise, are zero.

We transform each $\kappa_i$ into a sum of radicals. As mentioned at the end of Section 7.4 this can be done as follows. First compute the representation of $\kappa_i$ as a linear combination of powers of the primitive element $\eta_{it}$ for the extension $E_{it}$. Then using the same method as in Step 3 of the Algorithm Denesting Element transform this representation into a linear combination of the elements of the standard basis for this field.

The run time for this step is also covered by the previous run times since similar steps have already been applied in the denesting steps above, in particular, in Step 3 of the Algorithm Denesting Element. Also the coefficients in this representation are bounded in absolute value by $2^{O(L_n^2 N^2 \log N + dL \log N + K)}$. To prove this bound see Lemma 7.4.1 where a bound on the coefficients in the denesting of $\sqrt[\gamma]{\gamma}$ was given. Of course, the bound for the denesting also applies to the easier case we are presently considering. The estimate is even very crude since the term $dL \log N$, in the bound of Lemma 7.4.1 which comes from the bound on the denesting elements $\gamma^{(0)}$, will not show up in the bound for the coefficients $\kappa_i$. But it suffices for our purposes.

Observe that that each $\gamma_{it}$ is a linear combination over $\mathbb{Q}(\alpha)$ of elements of the form

$$\frac{1}{d_t \sqrt[\beta_{it}]{\theta_{it}}} \beta_{it}'.$$

$\beta_{it}, \beta_{it}'$ are elements of the standard basis of the radical extension $E_{it}$ and $\theta_{it} \in \mathbb{Q}(\alpha)$ is a denesting element for $d_t \sqrt[\beta_{it}]{\gamma_{it}^{d_t-1}} \beta_{it}$. By Corollary 7.3.2, p. 143, the coefficient size of $\theta_{it}$ is bounded by $2^{O(n^2 N^2 L + dL \log N + dK)}$.

In particular, $\gamma_{it}$ is a sum of real depth one radicals over $\mathbb{Q}(\alpha)$ containing at most $N$ terms. Hence, using the representation for $\kappa_i$ as a linear combination of the elements of the standard basis $E_{it}$ each product $\kappa_i \gamma_{it}$ is a sum of depth one radicals over $\mathbb{Q}(\alpha)$ and so is $S_t$.

The overall number of terms in the sums $S_t$, $t = 1, 2, \ldots, h$, is clearly bounded by $kN^2$ and, as mentioned above, the coefficients in these sums have coefficient size less than $2^{O(Ln^2 N^2 \log N + dL \log N + dK)}$.
To check whether \( S \) is zero we apply the method of Section 5. Hence we need to determine for ratios of the form

\[
\frac{\beta_{jt} \sqrt{\theta_{it} \beta_{jt}} - \beta_{it} \sqrt{\theta_{jt} \beta_{jt}}}{\beta_{it} \sqrt{\theta_{jt} \beta_{jt}}}
\]

whether they are elements of \( \mathbb{Q}(\alpha) \). Here \( \beta_{jt}, \beta_{it} \) are elements of the standard basis of \( E_{jt}, E_{it} \), respectively, appearing in the representation of \( \kappa_j, \kappa_i \) computed above. \( \theta_{it}, \theta_{jt}, \beta_{it}, \beta'_{it}, \beta_{jt}, \beta'_{jt} \) are as above.

As in the analysis of Step 1 and Step 3 of the previous section for these ratio tests we transform denominator and numerator separately into radicals \( \sqrt{\mu}, \sqrt{\nu} \), where \( D_{it}, D_{jt} \leq dN^2 \). This is possible since \( N^2 \) is an upper bound on the degree of the radical extension containing \( \kappa_i, \kappa_j, \gamma_i, \gamma_j \), and \( \gamma_t \), and the degree of any radical contained in this extension must divide the field degree. Analogously for the denominator.

One shows as in the analysis of Step 3 that

\[
\log[\mu], \log[\nu] \in \mathcal{O}(n^2 N^4 L + dN^2 L \log N + dN^2 K).
\]

For each of the \( \mathcal{O}(k^2 N^4) \) ratios these representations can be computed using \( \mathcal{O}(n^2 (\log^2 N + \log d)) \) operations on integers of size \( \mathcal{O}(n^2 N^4 L + dN^2 L \log N + dN^2 K) \). For example, if we denote for the sake of simplicity the degree of the extension generated by \( \kappa_i, \kappa_j, \gamma_i, \gamma_j \), and \( \gamma_t \) by \( N' \), \( N' \leq N^2 \), then \( D_{it} = d_i d_t N' \) and \( \beta_{jt} \) can be transformed into a radical of the form \( \sqrt[\rho_{ij}]{\mu} \), \( \rho \in \mathbb{Q}(\alpha) \) as follows.

Assume

\[
\beta_{jt} = \prod_{l=1}^{k} \sqrt[\rho_{ij}]{\rho_l^{\ell_{il}}},
\]

As in the analysis of Step 1 set \( \ell_i = N'/d_i \). Then

\[
\beta_{jt} = \sqrt[n']{\prod_{l=1}^{k} \rho_{ij}^{\ell_{il}}},
\]

To compute this representation we simply have to compute each \( \rho_l^{\ell_{il}} \), which can be done for a single \( \rho_l \) using \( \mathcal{O}(\log N') \) multiplications in \( \mathbb{Q}(\alpha) \) since \( \ell_i < d_i \ell_i < N' \). Since the number \( k \) of radicals \( \sqrt[\rho_{ij}]{\mu} \) is bounded by \( \log N' \) overall \( \mathcal{O}(\log^2 N') \) multiplications in \( \mathbb{Q}(\alpha) \) are needed. Then we multiply these powers. Taking the \( d_i d_t \)-th power of this element yields the representation \( \sqrt[\rho_{ij}]{\mu} \), \( \rho \in \mathbb{Q}(\alpha) \) for \( \beta_{jt} \). For the last step \( \mathcal{O}(\log d_i d_t) \) multiplications
are needed. By Lemma 5.5.5, p. 83, \([\rho] < 2^{\mathcal{L}dN^2 \log N}\). Now the time bounds follow from Lemma 5.5.2, p. 82.

Exactly the same analysis applies for the transformation of \(\beta'_{jt}\) into the appropriate form. To transform \(\sqrt{\theta_{it}}\beta_{it}\) into this form observe that we only need to transform \(\theta_{it}\) and \(\beta_{it}\) separately into \(N'\)-th roots. To do this the same process as above is applied. But observe that since \(\theta_{it}\) has coefficient size \(2^{\mathcal{O}(n^2 N^2 \mathcal{L} + dN \log N + dK)}\), the element \(\sqrt[172]{\theta'_{it}}\) in \(\theta_{it} = \sqrt[172]{\beta'_{it}}\) will have coefficient size \(2^{\mathcal{O}(n^2 N^4 \mathcal{L} + dN^2 \mathcal{L} \log N + dN^2 K)}\).

Multiplying the representations for \(\beta_{it}, \theta_{it}, \beta'_{jt}, \beta'_{jt}\) yields the desired result since this can be done by a constant number of multiplications.

Now Theorem 5.6.2, p. 90, or Table 1 on page 88 can be applied which shows that the last step in the General Denesting Algorithm requires

\[ \mathcal{O}(k^2 n^4 N^4 (n^2 N^2 \mathcal{L} + dN^2 \mathcal{L} \log N + dN^2 K)) \]

elementary operations on floating-point numbers and integers of size

\[ \mathcal{O}((n + d^2 N^4 + kN^2)(n^3 N^4 \mathcal{L} + dN^2 \mathcal{L} \log N + dN^2 K)), \]

since in this case \(k = kN^2, \max\{d_i\} \leq dN^2\), and \(L = n^2 N^4 \mathcal{L} + dN^2 \mathcal{L} \log N + dN^2 K\).

Except for the \(n^6\) factor these bounds are dominated by the time bounds for the denesting part.

The last claim of the theorem is an immediate consequence. □

The exact run time of the General Denesting Algorithm can be obtained from Table 2 by replacing \(K\) by \(dK\) in \(B\) and multiplying the number of operations by \(k^2\) and from Table 1 by replacing \(L\) by \(\mathcal{O}(\mathcal{L} n^2 N^2 \log N + d\mathcal{L} \log N + dK)\). The latter algorithm has to be applied at most \(k^4 N^4\) times.

Corollary 7.5.2 generalizes to the General Denesting Algorithm.

**Corollary 7.5.7** If the nested radicals \(\gamma_i, \kappa_i\) are as in Corollary 7.5.2 then the General Denesting Algorithm uses at most \(\mathcal{O}(k^2 d^5 m^3 n^7 N^{14} \mathcal{L})\) elementary operations on integers or floating-point numbers of size \(\mathcal{O}(kd^3 m^3 n^6 N^8 \mathcal{L})\).

In the very same way the corresponding result for complex radicals can be shown. The proof is slightly simpler since in the ratio tests we need not

\footnote{Observe that in the \(\mathcal{L}\) in Theorem 5.6.2 another factor of \(n\) is hidden.}
consider the ratio of nested radicals of different degree. Hence we need not
denest radicals whose degree is quadratic in the degree of the nested radicals
appearing in $S$.

**Theorem 7.5.8** Assume $\mathbb{Q}(\alpha)$ contains a primitive $d$-th root of unity. Sup-
pose $S = \sum_{i=1}^{k} \sqrt[d]{\gamma_i}$, is a depth two radical expression over $\mathbb{Q}(\alpha)$, such that
any depth one radical appearing in $S$ has the form $\sqrt[2]{p}$, $t|d$. Assume $\kappa_i, \gamma_i$ are
radical expressions satisfying the assumptions of Theorem 7.5.4. In partic-
ular their infinity norm is bounded by $2^K$ and for each expression $\gamma_i, \kappa_i$ an
integer $C_{\gamma_i}, C_{\kappa_i} < 2^K$ exists such that $C_{\kappa_i} \kappa_i$ or $C_{\gamma_i} \gamma_i$ is an algebraic integer.

Denote by $n$ the degree of the minimal polynomial $p$ of $\alpha$. Assume $l = \lceil \log |p|_2 \rceil$. $2^L$ is the maximum coefficient size of an element in $\mathbb{Q}(\alpha)$
appearing in the definition of $\kappa_i, \gamma_i$. Finally assume that for any fixed pair of
indices $i, j$ the radicals appearing in $\kappa_i, \kappa_j, \rho_i$, and $\rho_j$ generate an extension
of $\mathbb{Q}(\alpha)$ of degree at most $N$.

When applied to $S$ the General Denesting Algorithm determines in time polynomial in $k, n, l, L, N, d$, and the upper bound $K$ whether $S$ can be written
as a sum of real depth one radicals over $\mathbb{Q}(\alpha)$.

In particular, the General Denesting Algorithm uses at most $O(k^2 d^2 n^5 N^5 (L + dK) \log N)$ elementary operations on integers or floating-
point numbers of size $O(kd^2 n^5 N^4 (L + dK))$ plus the number of op-
erations needed to approximate each $\gamma_i$ with absolute error less than
$2^{-(46n^3 N^3 L + 12dN K)}$.

Within the same time bounds the algorithm decides whether $S = 0$.

The complex version of Corollary 7.5.2 generalizes to the General Den-
esting Algorithm.

**Corollary 7.5.9** If the nested radicals $\gamma_i, \kappa_i$ are as in Corollary 7.5.5 then
the General Denesting Algorithm uses at most $O(k^2 d^3 m^3 n^7 N^7 L \log N)$
elementary operations on integers or floating-point numbers of size $O(kd^3 m^3 n^5 N^6 \log N)$.

Before we finish this thesis let us mention some easy extensions of the
previous results to quotients of nested radicals. For the quotient of two
nested radicals we gave a denesting algorithm already in the General Den-
esting Algorithm.

If we are interested in one quotient of nested radicals of different degree
at least in the real case we can do slightly better than described in the
analysis of the General Denesting Algorithm. The improvement is based
on the following lemma which is an immediate consequence of Lemma 3.4, p. 24.

**Lemma 7.5.10** A ratio of real nested radicals \( \frac{\sqrt[\gamma_1]{d_1}}{\sqrt[\gamma_2]{d_2}} \) over a real field \( F \) denests using real radicals iff

- \( \sqrt[\gamma_1]{d_1} \) and \( \sqrt[\gamma_2]{d_2} \) denest using real radicals

and

- \( \sqrt[\gamma_1]{d_1} \sqrt[\gamma_2]{d_2} \) denests using real radicals.

Here \( d = \text{lcm}(d_1, d_2) \), \( d_i' = d_i/d \), and \( \gamma_i' \) denotes the denesting of \( \sqrt[\gamma_i]{d_i} \) if it exists.

Using this lemma, rather than applying the *Algorithm Denesting Element* to a nested radical of degree \( d_1d_2 \) we need to apply the algorithm only to nested radicals of degree \( \max\{d_1, d_2\} \).

For quotients of sums of nested radicals our results also yield an efficient solution. In fact, let \( \Sigma_1, \Sigma_2 \) be sums of real nested radicals of depth two. We may assume that the *General Denesting Algorithm* has already been applied to both sums. Hence no ratio of radicals appearing in the same sum denests. Suppose

\[
\frac{\Sigma_1}{\Sigma_2} = S,
\]

where \( S \) is a depth one expression. This implies

\[
\Sigma_1 - S\Sigma_2 = 0.
\]

Applying to this sum of nested radicals Theorem 6.1.2, p. 100, or its complex equivalent, Theorem 6.1.5, p. 104, shows that for any nested radical \( \sqrt[\gamma]{\eta} \) in \( \Sigma_1 \) there exists a *unique* nested radical \( \sqrt[\eta]{\tau} \) in \( \Sigma_2 \) such that their ratio denests.

Using the previous results we may compute one such denesting, say,

\[
\frac{\sqrt[\gamma]{\eta}}{\sqrt[\eta]{\tau}} = \theta.
\]

If the coefficients of \( \sqrt[\gamma]{\eta} \) and \( \sqrt[\eta]{\tau} \) in \( \Sigma_1, \Sigma_2 \) are \( \kappa, \mu \), respectively, then \( S \) is already uniquely determined as

\[
S = \frac{\theta\kappa}{\mu}.
\]

174
Finally we check whether $S$ satisfies $\Sigma_1 - S \Sigma_2 = 0$ using the *General Denesting Algorithm*.

Obviously the run time is determined by the last step so the analysis of Theorem 7.5.6 applies.

The next step is to go to sums of ratios of sums of nested radicals. Here we know nothing better than just transforming the problem into a single ratio of sums of nested radicals. But the sums in the numerator and denominator may have length exponential in the length of the original sum. On the other hand, if the bounds from Section 6 are tight as is to be expected then the output size of a denesting may also be exponential in the length of the sum of ratios.

But recall that in algorithmic algebra we usually assume that numbers are given in some basis representation, that is in our case, as a linear combination of nested radicals.
Appendix: Roots of Unity in Radical Extensions of the Rational Numbers

The goal of this appendix is to give the proof of

**Theorem 6.4.4** Let $F = \mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k})$ be a real radical extension of $\mathbb{Q}$. If $\zeta_m$ is a primitive $m$-th root of unity then $F(\zeta_m)$ contains at most $24m$ different roots of unity. Moreover, the constant $24$ is best possible, i.e., there are real radical extensions $F$ of $\mathbb{Q}$ and $m \in \mathbb{N}$ such that $F(\zeta_m)$ contains exactly $24m$ different roots of unity.

**Proof:** First we reformulate the problem a bit.

**Lemma A 1** The number $M$ of roots of unity in $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ is the maximum of all numbers $N$ such that $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ contains a primitive $N$-th root of unity. Moreover, $m$ divides $M$.

**Proof:** Assume the field contains no primitive $M$-th root of unity, instead assume $N < M$ is the largest number such that $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ contains an $N$-th primitive root of unity. Then $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ contains a primitive $N$-th root of unity and a primitive $K$-th root of unity $\zeta_{KN}$ for $(N, K) = 1, K > 1$. By Lemma 3.7, p. 26, in this case $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ also contains a primitive $K$-th root of unity. This contradicts the maximality of $N$, so $M = N$.

But then all roots of unity in $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ must be a power of $\zeta_M$. In particular, the primitive $m$-th roots of unity $\zeta_m$ must be a power of $\zeta_M$. This is possible if and only if $m$ divides $M$. \[\square\]

In view of these facts we can reformulate the original problem. We have to determine the largest multiple $M$ of $m$ such that $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m) = \mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_M)$ for primitive $m$-th and $M$-th roots of unity $\zeta_m, \zeta_M$. Moreover, instead of answering this question for the field $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$ we will answer it for $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_{m'})$ where $m' = \text{lcm}(4, m)$. The number $M$ deduced in this way will clearly be an upper bound on the number of roots of unity in $\mathbb{Q}(\sqrt[q_1]{d_1}, \ldots, \sqrt[q_k]{d_k}, \zeta_m)$.

Now assume that the prime factorization of $m'$ is given by $m' = 2^e \prod_{i=1}^{e_i} p_i^{e_i}$, $p_i$ prime, $e, e_i \in \mathbb{N}, e \geq 2$. Then $M$ can be written as $M = 2^{2e} \prod_{i=1}^{e'} q_i^{f_i} \prod_{i=1}^{e'} p_i^{e_i}$, $q_i$ prime, $e', f_i \in \mathbb{N}, e' \geq e$, where the $q_i$'s need not be distinct from the $p_i$'s.
If $\mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_{m'}) = \mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_M)$ then the first field must be the same as $\mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_{M_i})$, for all $i = 0, 1, \ldots, l'$, where $M_0 = 2^{e'} \prod_{i=1}^l p_i^{e_i}$ and $M_i = q_i^{f_i} 2^{e'} \prod_{i=1}^l p_i^{e_i}$ for $i > 0$. We will show that this is possible only for $e' - e = 1$ and $q_i^f = 3$. This implies $M \leq 6m' \leq 24m$ and will therefore prove the upper bound of the theorem.

To prove the claim we consider for each $M_i$, $i = 0, 1, \ldots, l'$ a field $E_i$ such that if $\mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_{m'}) = \mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_{M_i})$ then $E_i$ must be a subfield of $E_i$. Clearly the degrees of $\mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_{M_i})$ and of $\mathbb{Q}(\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}, \zeta_{M_i})$ over $E_i$ have to be the same. From this property the claim will easily follow.

We will choose the field $E_i$ to be the field generated by the real radicals $\sqrt[4]{q_1}, \ldots, \sqrt[4]{q_k}$ and all real square roots in $\mathbb{Q}(\zeta_{M_i})$.

**Lemma A 2** Let $m \in \mathbb{N}$ such that $4|m$. If $m = 2^e \prod_{i=1}^l p_i^{e_i}$, $e_i \geq 1$, $e \geq 2$, is the prime factorization of $m$ then the subfield of $\mathbb{Q}(\zeta_m)$ generated by all real square roots in $\mathbb{Q}(\zeta_m)$ is $\mathbb{Q}(\sqrt[4]{p_1}, \ldots, \sqrt[4]{p_l})$ if $e = 2$ and $\mathbb{Q}(\sqrt[4]{p_1}, \ldots, \sqrt[4]{p_l}, \sqrt{2})$ if $e > 2$.

**Proof:** First all quadratic subfields of $\mathbb{Q}(\zeta_m)$ will be determined. By Galois theory these subfields correspond to subgroups of the Galois group of $\mathbb{Q}(\zeta_m)$ over $\mathbb{Q}$ of order $\varphi(m)$, $\varphi(m) = [\mathbb{Q}(\zeta_m) : \mathbb{Q}]$. As we mentioned above (see Lemma 2.3, p. 16) the Galois group of this extension is isomorphic to $\mathbb{Z}^*_m$, the multiplicative group of integers taken modulo $m$ between 1 and $m$ which are relatively prime to $m$. Hence it is abelian. By the following result due to G. Birkhoff [Bi] the number of quadratic subfields of $\mathbb{Q}(\zeta_m)$ equals the number of subgroups of $\mathbb{Z}^*_m$ of order 2.

**Lemma A 3 (Birkhoff)** If $G$ is an abelian group of order $n$ then the number of subgroups of order $\frac{n}{d}$, $d|n$, equals the number of subgroups of order $d$.

As is well-known from number theory (see [IR]) $\mathbb{Z}^*_m$ can be written as a direct product

$$\mathbb{Z}^*_m = \mathbb{Z}^*_{2e} \times \mathbb{Z}^*_{p_1^{e_1}} \times \mathbb{Z}^*_{p_2^{e_2}} \times \cdots \times \mathbb{Z}^*_{p_l^{e_l}},$$

where $\mathbb{Z}^*_p$ is a cyclic group of order $p^{e_i-1}(p_i - 1)$ and $\mathbb{Z}^*_{2e}$ is either a cyclic group of order 2 or a direct product of two cyclic groups $C_1, C_2$, one of order 2 and the other of order $2^{e-2}$.

Each subgroup of order 2 of $\mathbb{Z}^*_m$ must be cyclic. Hence we have to determine all elements in $\mathbb{Z}^*_m$ of order 2. By the above representation for $\mathbb{Z}^*_m$ these elements correspond to products $h_1h_2g_1 \cdots g_l$, where $h_1 \in C_1, h_2 \in C_2, $$g_1, \ldots, g_l$
$C_2, g_i \in \mathbb{Z}_{p_i}^*$, and each element is either the unit element of that group or an element of order 2. If $e = 2$ then we have to dismiss the second factor.

By group theory any cyclic group of order $d$ contains for each divisor $d'$ of $d$ exactly one element of order $d'$. Hence there are $2^{l+1} - 1$ or $2^{l+2} - 1$ elements of order 2 in $\mathbb{Z}_m^*$ depending on whether $e = 2$ or $e > 2$. The $-1$-terms occurs because we are not allowed to take the unit element from each subgroup. Accordingly, $\mathbb{Q}(\zeta_m)$ has either $2^{l+1} - 1$ or $2^{l+2} - 1$ quadratic subfields.

Next observe that $\mathbb{Q}(\zeta_{p_i}), \mathbb{Q}(\zeta_4)$ are subfields of $\mathbb{Q}(\zeta_m)$. And if $8 | m$ then $\mathbb{Q}(\zeta_8)$ is also a subfield. A well-known result in algebraic number theory (see for example [Ja]) states that the unique quadratic subfield of $\mathbb{Q}(\zeta_{p_i})$ is generated by $\sqrt{(-1)p_i}$ if $p_i \equiv 3 \mod 4$ and is generated by $\sqrt{p_i}$ if $p_i \equiv 1 \mod 4$. Moreover, $\mathbb{Q}(\zeta_4)$ is of course generated by $\sqrt{-1}$ and $\mathbb{Q}(\zeta_8)$ has three quadratic subfields generated by $\sqrt{-1}, \sqrt{2},$ and by $\sqrt{-2}$.

Therefore $\mathbb{Q}(\zeta_m)$ contains

$$\sqrt{(-1)f_1^{2}f_2^{l+2}p_1^{f_3} \cdots p_l^{f_{l+2}}},$$

where each $f_i$ is either 0 or 1 and in case $e = 2$ $f_2$ is always 0.

As is easily seen (use Theorem 3.9, p. 28, for example, but it can also be proven directly) these square roots generate pairwise distinct quadratic subfields of $\mathbb{Q}(\zeta_M)$. Since this yields $2^{l+1} - 1$ or $2^{l+2} - 1$ distinct quadratic fields depending on whether $e = 2$ or $e > 2$ these must be all quadratic subfields. Hence a real square root that is contained $\mathbb{Q}(\zeta_m)$ must generate one of the fields

$$\mathbb{Q}\left(\sqrt{2^{f_2}f_1^{f_3} \cdots p_l^{f_{l+2}}} \right), f_1 = 0, 1, f_2 = 0 \text{ if } e = 2.$$ 

Since all these fields are contained in $\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_l})$ if $e = 2$ and in $\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_l}, \sqrt{2})$ if $e > 2$ the lemma follows. 

Denote the field generated by the real square roots contained in $\mathbb{Q}(\zeta_M)$ and by the real radicals $\sqrt{\sqrt{p_1}}, \ldots, \sqrt{\sqrt{p_l}}$ by $E_i$. Hence $E_i \subset \mathbb{Q}(\zeta_M)$ and $\mathbb{Q}(\sqrt{\sqrt{p_1}}, \ldots, \sqrt{\sqrt{p_l}}, \zeta_M) = E_i(\zeta_M)$. Moreover, if $\zeta_M \in \mathbb{Q}(\sqrt{\sqrt{p_1}}, \ldots, \sqrt{\sqrt{p_l}}, \zeta_m')$ then $\mathbb{Q}(\sqrt{\sqrt{p_1}}, \ldots, \sqrt{\sqrt{p_l}}, \zeta_m') = E_i(\zeta_m')$ and therefore $E_i(\zeta_m') = E_i(\zeta_M)$. In particular, the degree of $E_i(\zeta_m')$ over $E_i$ must be equal to the degree of $E_i(\zeta_M)$ over $E_i$. 

178
Applying Theorem 2.4, p. 16, to \( K = \mathbb{Q}, E = \mathbb{Q}(\zeta_m), F = E_i \) and to \( K = \mathbb{Q}, E = \mathbb{Q}(\zeta_{m'}), F = E_i \) shows that this implies

\[
[\mathbb{Q}(\zeta_{m'}) : \mathbb{Q}(\zeta_m) \cap E_i] = [\mathbb{Q}(\zeta_m) : \mathbb{Q}(\zeta_{m'}) \cap E_i], i = 0, 1, \ldots, l'.
\]

Next we determine how the intersections look like.

**Lemma A 4** Let \( \sqrt[p_1]{q_1}, \ldots, \sqrt[p_k]{q_k} \) be real radicals and \( \zeta_m \) be a primitive \( m \)-th root of unity. By \( E \) denote the subfield of \( \mathbb{Q}(\sqrt[p_1]{q_1}, \ldots, \sqrt[p_k]{q_k}, \zeta_m) \) that is generated by the radicals \( \sqrt[p_i]{q_i} \) and by the real square roots contained in \( \mathbb{Q}(\zeta_m) \). Then \( E \cap \mathbb{Q}(\zeta_m) \) is the field generated by the real square roots in \( \mathbb{Q}(\zeta_m) \).

**Proof:** Since \( E \cap \mathbb{Q}(\zeta_m) \) is a subfield of the real radical extension \( E \) it must be generated by real radicals (see Theorem 3.13, p. 32).

Since \( E \cap \mathbb{Q}(\zeta_m) \) is a real radical extension contained in \( \mathbb{Q}(\zeta_m) \) it must be generated by square roots (see Lemma 3.12, p. 32). Also by Lemma 3.12, p. 32, the field generated by all real square roots in \( \mathbb{Q}(\zeta_m) \) is the largest possible subfield of \( \mathbb{Q}(\zeta_m) \) generated by real radicals.

By definition of \( E \) this field is also a subfield of \( E \). The lemma follows. \( \square \)

Combining Lemma A 2 and Lemma A 4 shows

- If \( e > 2 \) then
  \[
  F_i = \mathbb{Q}(\zeta_m) \cap E_i = \mathbb{Q}(\sqrt[2]{2}, \sqrt[p_1]{q_1}, \ldots, \sqrt[p_l]{q_l}), i = 1, 2, \ldots, l',
  \]
  and
  \[
  F = \mathbb{Q}(\zeta_{m'}) \cap E_i = F_0 = \mathbb{Q}(\zeta_{m_0}) \cap E_i = \mathbb{Q}(\sqrt[2]{2}, \sqrt[p_1]{p_1}, \ldots, \sqrt[p_l]{p_l}).
  \]

- If \( e = e' = 2 \) then
  \[
  F_i = \mathbb{Q}(\zeta_m) \cap E_i = \mathbb{Q}(\sqrt[p_1]{p_1}, \ldots, \sqrt[p_l]{p_l}, \sqrt[q_i]{q_i}), i = 1, 2, \ldots, l',
  \]
  and
  \[
  F = \mathbb{Q}(\zeta_{m'}) \cap E_i = F_0 = \mathbb{Q}(\zeta_{m_0}) \cap E_i = \mathbb{Q}(\sqrt[p_1]{p_1}, \ldots, \sqrt[p_l]{p_l}) \text{ for all } i.
  \]
• If $e = 2$, $e' > 2$ then

$$F_i = Q(\zeta_{M_i}) \cap E_i = Q(\sqrt{p_1}, \ldots, \sqrt{p_l}, \sqrt{q_i}), \ i = 1, 2, \ldots, l', \quad F_0 = Q(\zeta_{M_0}) \cap E_0 = Q(\sqrt{2}, \sqrt{p_1}, \ldots, \sqrt{p_l}),$$

and

$$F = Q(\zeta_{M'}) \cap E_i = Q(\sqrt{p_1}, \ldots, \sqrt{p_l}) \text{ for all } i.$$

Since field degrees are multiplicative

$$\frac{\varphi(M_i)}{\varphi(m')} = \frac{[Q(\zeta_{M_i}) : Q]}{[Q(\zeta_{M'}) : Q]} = \frac{[F_i : Q]}{[F : Q]}.$$

First consider $i = 0$ and assume $e' > e$. In this case

$$\frac{\varphi(M_0)}{\varphi(m')} = 2^{e'-e}$$

but

$$\frac{[F_0 : Q]}{[F : Q]} = 2$$

if $e = 2$. Otherwise this ratio is 1. Hence if $e = 2$ then $e'$ can be at most 3 and if $e > 2$ then $e = e'$.

For $i > 0$ we can use a similar argument.

$$\frac{\varphi(M_i)}{\varphi(m')} = q_i^{f_i-1}(q_i - 1)$$

if $q_i$ is distinct from all $p_j$'s. Otherwise

$$\frac{\varphi(M_i)}{\varphi(m')} = q_i^{f_i}.$$

On the other hand

$$\frac{[F^{(i)} : Q]}{[F : Q]} = 2$$

or

$$\frac{[F^{(i)} : Q]}{[F : Q]} = 1$$

depending on whether $q_i$ is distinct from the $p_j$'s or not.
Hence \(q_i^{f_i-1}(q_i - 1) = 2\) or \(q_i^{f_i} = 2\). The second case is impossible for an odd prime and the first one is possible if and only if \(q_i = 3\) and \(f_i = 1\). As mentioned this proves the upper bound.

It remains to show that this bound is optimal. To do so let \(m\) be a positive integer such that \(\gcd(24, m) = 1\). Moreover let \(m\) be divisible by a prime \(p\) satisfying \(p \equiv 3 \mod 4\). Consider \(Q(\sqrt{2}, \sqrt{3}, \sqrt{p}, \zeta_m)\), where \(\zeta_m\) is a primitive \(m\)-th root of unity.

As noted above \(Q(\zeta_m)\) contains \(\sqrt{-p}\). Hence \(\sqrt{-1} \in Q(\sqrt{2}, \sqrt{3}, \sqrt{p}, \zeta_m)\). Therefore this field contains
\[
\frac{1}{\sqrt{2}}(1 + \sqrt{-1}) \text{ and } \frac{1}{2}(1 + \sqrt{-3}).
\]
the first number is a primitive 8-th root of unity and the second one a primitive 3-rd root of unity. By Lemma 3.7, p. 26, this implies that \(Q(\sqrt{2}, \sqrt{3}, \sqrt{p}, \zeta_m)\) contains a 24\(m\)-th primitive root of unity.

The following corollary is an immediate consequence of the previous theorem.

**Corollary A 5** Let \(m \in \mathbb{N}\). Both numbers \(\sin \frac{2\pi m}{m}\) and \(\cos \frac{2\pi m}{m}\) can be written as linear combination of real radicals over \(Q\) if and only if \(m|24\).
References


Summary

In this thesis we describe several simplification algorithms for expressions involving radicals, for example sums of square roots. We show how to transform any sum of square roots into a sum of linearly independent square roots. After this transformation it is very easy to determine whether the sum is zero.

More generally, it is shown how to determine for any sum of real radicals over the rational numbers in time polynomial in the input size of the sum whether it is zero. This contrasts to the fact, that until now no efficient algorithms to determine the sign of a sum of radicals are known.

Other examples of radical expressions that are simplified are the so-called nested radicals. The problem of denesting nested radicals is best explained by the following examples which can be found in the notebook of the Indian mathematician Ramanujan.

\[
\sqrt{5} - \sqrt{4} = \frac{1}{3} \left( \sqrt{2} + \sqrt{20} - \sqrt{25} \right)
\]

\[
\sqrt[3]{7} \sqrt[3]{20} - 19 = \sqrt[3]{5} - \sqrt[3]{2} \cdot \sqrt[3]{3}.
\]

The expressions on the left-hand side of the equations have nesting depth two while the expressions on the right-hand side have nesting depth one. In the thesis it is shown that for a large class of nested radicals of depth two a denesting can be found in polynomial time. This class contains all of Ramanujan’s examples. Although in many respects more general, the algorithms known so far for denesting radicals cannot handle the denestings found by Ramanujan.

From a theoretical point of view the results mentioned above are based on the fact that real radicals over a real field \( F \) are already linearly independent if any two of them are linearly independent. The proof of this fact in turn is based on a theorem due to C. L. Siegel that determines to some extend the structure of real radical extensions. We give a simplified proof of this theorem.

From an algorithmic point of view the results are based on an algorithm that, given an algebraic number field, determines in polynomial time for an element in this field its exact representation provided an upper bound on the
representation size of the element and an approximation to this element are given. The main ingredient to this algorithm is the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász.
Zusammenfassung


Andere Beispiele von Ausdrücken, die vereinfacht werden, sind geschachtelte Wurzelausdrücke. Das Problem wird am besten durch die folgenden den Notizbüchern des indischen Mathematikers Ramanujan entnommenen Beispiele demonstriert.

\[
\sqrt{3} \sqrt[3]{5} - 3 \sqrt[3]{4} = 1
\]

\[
\frac{3}{\sqrt[6]{7} \sqrt[3]{20} - 19} = \sqrt[3]{5} - \frac{1}{3}. \]


In algorithmischer Hinsicht beruhen die Ergebnisse der Dissertation auf einem Algorithmus, der, gegeben einen algebraischen Zahlkörper, in poly-
Johannes Blömer

Lebenslauf

23. Mai 1964 geboren in Dinklage

1970 - 1974 Besuch der Grundschule Dinklage

1974 - 1983 Besuch des Gymnasiums Lohne

Mai 1983 Abitur am Gymnasium Lohne

Okt.1983 - Feb. 1989 Studium der Mathematik, Geschichte und Philosophie an der FU Berlin

Juli 1985 Vordiplom in Mathematik

Feb. 1989 Diplom in Mathematik

April 1989 - Mai 1989 Werkvertrag an der FU Berlin

seit Mai 1989 wissenschaftlicher Mitarbeiter im DFG-Projekt “Georechner” von Prof. Dr. H. Alt

seit Oktober 1991 wissenschaftlicher Mitarbeiter an der FU Berlin