CALCULATION OF FUNDAMENTAL UNITS IN SOME TYPES OF QUARTIC NUMBER FIELDS

A thesis presented in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

in Mathematics

at Massey University

Neville Stuart Jeans

1984

ABSTRACT

Dirichlet's theorem describing the structure of the unit group of the ring of integers of an algebraic number field shows that the units are generated by a primitive root of unity of the field plus a finite set of units called a fundamental system of units. However Dirichlet's theorem does not suggest any method by which a fundamental system of units can be obtained. In this thesis we consider the problem of calculating a fundamental system of units for certain types of quartic field which are a quadratic extension of a quadratic field $Q(\delta)$. Our attention is mainly centered on type I quartic fields, that is quartic fields for which $Q(\delta)$ is complex. In such cases a fundamental system of units contains a single unit called a fundamental unit.

To calculate fundamental units of type I quartic fields we use the simple continued fraction algorithm, real quadratic field case as a guide. This topic is reviewed in chapter one where we also note Voronoi's view of simple continued fractions in terms of relative minima of a Z module.

In chapter two we consider the idea of relative minima of a module over a ring of complex quadratic integers. Basically we generalize the simple continued fraction algorithm which calculates best approximations to a real number using rational integer coefficients to an algorithm which calculates best approximations to a complex number using complex quadratic integer coefficients. The ideas are developed with respect to an arbitrary complex quadratic field $Q(\delta)$ and show many similarities to the simple continued fraction algorithm. (Existing work of this nature restricts its attention to cases where $Q(\delta)$ has class number one). We obtain an algorithm which is periodic for complex numbers w satisfying

$w^2 \in Q(\delta)$, $w \notin Q(\delta)$. This enables us to calculate units of type I quartic fields.

In chapter three we consider quartic fields $Q(\Gamma)$ which are a quadratic extension of a quadratic field $Q(\delta)$. In section one we express the ring of integers of $Q(\Gamma)$ in terms of the integers of $Q(\delta)$ thereby recognising four forms which these rings may take. In section two we consider the problem of calculating fundamental units of type I quartic fields. The algorithm developed in chapter two is only guaranteed to locate a fundamental unit when the ring of integers of $Q(\Gamma)$ is of the simplest of the four forms mentioned above. A modified version of the algorithm allows us to calculate a fundamental unit when the ring of integers of $Q(\Gamma)$ is of the second simplest form. For the two remaining forms we obtain a unit U which may or may not be fundamental. We therefore develop an algorithm which calculates a fundamental unit from U. To illustrate the use of our algorithms we calculate fundamental units for the type I quartic fields

$$Q(\sqrt[4]{D}),\ D \in Z,\ -99 \leq D \leq -1$$

Finally in section three we consider the calculation of a fundamental system of units for type IIb quartic fields, that is semi-real quartic fields which are a quadratic extension of a real quadratic field. A connection between type IIb and type I quartic fields enables us to calculate fundamental systems of units for type IIb quartic fields.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr M.D. Hendy and Dr K.L. Teo for the advice and encouragement they have offered during the preparation of this thesis. Thanks also to Gail Tyson for typing this thesis.

CONTENTS

REFERENCE PAGES FOR NOTATION AND TERMINOLOGY

The list which follows is not exhaustive. We have omitted most standard notation and terminology plus most notation and terminology which has a localised usage.

CHAPTER ONE

ALGEBRAIC NUMBER FIELDS, QUADRATIC FIELDS, AND SIMPLE CONTINUED FRACTIONS

This introductory chapter is intended to serve two main purposes. Firstly to introduce the notation related to the theories of algebraic number fields, quadratic fields, and simple continued fractions which is to be used in this thesis. Secondly to record for the purposes of easy reference and brief review the main ideas, results and algorithms from these theories which are either used or referred to in chapters two and three. Note that the presentation in this chapter is intended for the reader who is already familiar with the terminology and ideas of these theories. For a detailed coverage of the material noted in this chapter we will refer the reader to an appropriate selection of texts from the literature.

SECTION ONE

ALGEBRAIC NUMBER FIELDS, MODULES, AND MISCELLANEA

The symbols defined below will have the same meaning throughout this thesis.

$\mathbb{Z}^+$ - the set of positive rational integers

$\mathbb{Z}$ - the set of rational integers

$\mathbb{Q}$ - the set of rational numbers

$\mathbb{R}^+$ - the set of positive real numbers

$\mathbf{R}$ - the set of real numbers

$\mathbf{C}$ - the set of complex numbers

i - the square root of -1

$\mathrm{Re}(x)$ - the real part of $x \in \mathbf{C}$

$\mathrm{Im}(x)$ - the imaginary part of $x \in \mathbf{C}$

$\arg(x)$ - the argument of $x \in \mathbf{C}$

$\bar{x}$ - the complex conjugate of $x \in \mathbf{C}$

$\sqrt{x}$ - the square root of $x \in \mathbf{C}$ having $\arg(\sqrt{x}) \in (-\frac{\pi}{2}, \frac{\pi}{2}]$

$R\backslash S$ - the set of elements contained in set R but not in set S

$[x]$ - the greatest integer function for $x \in \mathbf{R}$

We shall also use the notation $(a_1,a_2,\ldots,a_k)$ to denote the greatest common divisor of $a_1, a_2,\ldots, a_k \in Z$, and $\{r\}$ to denote the nearest integer function for $r \in \mathbf{R}$, that is

$$\{r\} = a \in Z \text{ where } |r-a| \leq |r-b| \ \forall\ b \in Z \qquad (1)$$

(Ties to be settled by some arbitrary rule). Note that the last two notations are not exclusively reserved for the two purposes indicated. For example ordered pairs will be denoted by (α,β) and certain infinite sequences will be denoted by $\{A_k\}$. However the meaning will always be clear from the context. All other notation used in this thesis will be introduced as it is required.

The general theory and terminology for the subject of algebraic number fields should be well known to any reader of this thesis. Texts such as Adams and Goldstein [1976], Borevich and Shafarevich [1966], Cohn [1962], Richman [1971], and Stewart and Tall [1979] amongst many others cover to varying degrees the relevant general background.

However at this point we briefly note a number of basic terms and results from this subject area in order to define notation used in this thesis.

Let α be an algebraic number of degree n over Q. Then we use $Q(\alpha)$ to denote the algebraic number field of degree n over Q formed by adjoining α to Q. The degree of this extension field is symbolized by writing $[Q(\alpha) : Q] = n$. The ring of integers of $Q(\alpha)$ will be denoted by $Z(\alpha)$.

Now suppose β is of degree m over $Q(\alpha)$. Then we use $Q(\alpha)(\beta)$ to denote the algebraic number field formed by adjoining β to $Q(\alpha)$ and $Z(\alpha)(\beta)$ to denote the ring of integers of $Q(\alpha)(\beta)$. We write

$$[Q(\alpha)(\beta) : Q(\alpha)] = m$$

$$[Q(\alpha)(\beta) : Q] = [Q(\alpha)(\beta) : Q(\alpha)][Q(\alpha) : Q] = mn$$

to symbolise the facts that $Q(\alpha)(\beta)$ is a degree m extension of $Q(\alpha)$ and a degree of mn extension of Q. We can always find $\gamma \in Q(\alpha)(\beta)$ of degree mn over Q such that $Q(\gamma) = Q(\alpha)(\beta)$. (Thus $Z(\gamma) = Z(\alpha)(\beta)$).

The norm function from $Q(\alpha)$ to Q will be denoted by N while the relative norm function from $Q(\alpha)(\beta)$ to $Q(\alpha)$ will be denoted by N_α.

Of particular interest in this thesis is the structure of the unit group of $Q(\alpha)$ (more correctly of $Z(\alpha)$) which is described by the following theorem. Note that by a real field we mean a field F such that $F \subseteq \mathbb{R}$ and by a non-real field we mean a field F such that $F \subseteq \mathbb{C}$ but $F \not\subseteq \mathbb{R}$.

THEOREM 1.1 Dirichlet's Theorem (Stewart and Tall [1979,p219])

Suppose α (equivalently $Q(\alpha)$) has s real conjugates (s real conjugate fields) and 2t non-real conjugates (2t non-real conjugate fields) where $n = s + 2t$ is the degree of the minimal polynomial for α. Let ξ be a

primitive u^{th} root of unity which generates the (finitely many) roots of unity of $Z(\alpha)$. Then we can find units

$$U_1, U_2, \ldots, U_r \in Z(\alpha), \quad r = s + t - 1$$

such that the units of $Z(\alpha)$ are precisely those numbers of the form

$$\xi^a U_1^{a_1} U_2^{a_2} \ldots U_r^{a_r}, \quad 1 \leq a \leq u, \; a_j \in Z, \; j = 1,2,\ldots,r \qquad //$$

The set $\{U_1, U_2, \ldots, U_r\}$ (which is not unique) will be referred to as a fundamental system of units of $Q(\alpha)$ and can be characterised as follows. Let $\sigma_1, \sigma_2, \ldots, \sigma_s$ denote the real monomorphisms of $Q(\alpha)$ into $\mathbb{C}$ and let $\sigma_{s+1}, \bar{\sigma}_{s+1}, \ldots, \sigma_{s+t}, \bar{\sigma}_{s+t}$ denote the pairs of complex monomorphisms of $Q(\alpha)$ into $\mathbb{C}$. Then

$\{U_1, U_2, \ldots, U_r\}$ is a fundamental system of units of $Q(\alpha)$

if and only if

$U_1, U_2, \ldots, U_r$ are units of $Q(\alpha)$ for which the determinant of any order r square submatrix of

$$\begin{bmatrix} b_1 \ell n|\sigma_1(U_1)| & \cdots & b_{s+t} \ell n|\sigma_{s+t}(U_1)| \\ \vdots & & \vdots \\ b_1 \ell n|\sigma_1(U_r)| & \cdots & b_{s+t} \ell n|\sigma_{s+t}(U_r)| \end{bmatrix} \qquad b_j = \begin{cases} 1 \text{ if } j \leq s \\ 2 \text{ if } j > s \end{cases}$$

is non-zero and of minimal magnitude R amongst all sets of r units of $Q(\alpha)$. (The columns of the above matrix sum to zero so R does not depend on which column is deleted to give the order r submatrix).

The minimal magnitude value R is called the regulator of $Q(\alpha)$.

Note that if $s > 0$ then $u = 2$, $\xi = -1$. If $r = 1$ then a fundamental system of units contains just one unit which is normally referred to as a fundamental unit.

Dirichlet's theorem only asserts the existence of a fundamental system of units and does not suggest any way by which such a system might be calculated. One of the principal aims of this thesis is to develop algorithms for calculating fundamental systems of units for certain types of quartic fields.

In this thesis we will make extensive use of modules of the form

$$R[A_1, A_2, \ldots, A_k] = \{\alpha_1 A_1 + \alpha_2 A_2 + \ldots + \alpha_k A_k : \alpha_j \in R\}$$

where R is the ring of integers of an algebraic number field (that is $R = Z$ or $R = Z(\alpha)$) and the generators $A_1, A_2, \ldots, A_k$ are complex numbers or two dimensional complex vectors. Such modules will be referred to as R modules and a minimal set of generators will be called a basis. A basis for an R module is not unique but two bases are related by an R integral matrix whose determinant is a unit of R.

We finish this section by noting a result which will be used in chapter two. Suppose we have a collection of non-overlapping circles (that is the area of intersection of any two distinct circles is zero) of radius r in the plane. For $t \in \mathbf{R}^+$ let n(t) denote the number of these circles which lie completely within a circle of radius t centered on the origin. The packing constant for this collection of circles is defined to be

$$\lim_{t\to\infty} \left(\frac{\text{area of } n(t) \text{ circles}}{\text{area of circle of radius } t}\right) = \lim_{t\to\infty} \frac{n(t)r^2}{t^2}$$

The densest possible packing of equal radius circles in the plane is hexagonal and this produces the largest possible packing constant of $\pi/2\sqrt{3} \approx .907$. (See Fejes Tóth [1953, chapter III]).

SECTION TWO

QUADRATIC FIELDS

In this section we define the notation related to quadratic fields which will be used in this thesis. For ease of reference we also include a number of the more important results and ideas which are relevant to later chapters. A major part of this thesis will deal with results related to quartic fields which are quadratic extensions of the quadratic fields described in this section. Consequently the case discussed below will occasionally be referred to as the "standard quadratic case" in later chapters.

The literature of course contains numerous works which deal with the subject of quadratic fields. For example Cohn [1962], Adams and Goldstein [1976], Hardy and Wright [1979] and Richman [1971] to name but a few. We have mainly used Cohn as a source for the following results.

Let d be a square-free rational integer, $d \neq 1$, and let

$$\delta = \sqrt{d}$$

Then $Q(\delta)$ is a quadratic field. Where necessary we distinguish the two subcases corresponding to $d > 0$ and $d < 0$ by using the prefixes "real" and "complex". The integers of the field $Q(\delta)$ are referred to as quadratic integers and the set of all integers of $Q(\delta)$ is of course denoted by $Z(\delta)$. Define

$$c = \begin{cases} 2 & \text{if } d \equiv 1 \pmod 4 \\ 1 & \text{otherwise} \end{cases} \qquad (2)$$

Then

$$Z(\delta) = \{ (a+b\delta)/c : a,b \in Z, \quad a \equiv b \pmod c \}$$

Note that whenever we use the notation $(a+b\delta)/c$ to denote a $Z(\delta)$ integer

then we will assume that the above conditions on a,b,c are understood. In particular c will be used exclusively throughout this thesis to denote the value defined in (2). If we define

$$\omega = \begin{cases} (1+\delta)/2 & \text{if } d \equiv 1 \pmod 4 \\ \delta & \text{otherwise} \end{cases}$$

then we have the alternative form

$$Z(\delta) = \{a+b\omega : a,b \in Z\} = Z[1,\omega]$$

for the integers of $Q(\delta)$. We shall make use of both forms in this thesis. Note that the second form implies that $\{1,\omega\}$ is an integral basis for $Q(\delta)$ and so we have that the discriminant of the field $Q(\delta)$ is

$$\Delta = \begin{cases} d & \text{if } d \equiv 1 \pmod 4 \\ 4d & \text{otherwise} \end{cases}$$

For $\alpha = a + b\delta \in Q(\delta)$ $(a,b \in Q)$ we use α' to denote the conjugate $a - b\delta$ of α. The norm function from $Q(\delta)$ to Q is of course $N(\alpha) = \alpha\alpha' = a^2 - b^2 d$.

The following result will prove useful on several occasions in later chapters.

THEOREM 1.2 (Cohn [1962,p128, "Hurwitz's Lemma"])

If $\alpha,\beta \in Z(\delta)$, $g \in Z$ and $g \mid (N(\alpha), N(\beta), N(\alpha+\beta))$ then

$$g \mid \alpha\beta', \alpha'\beta \qquad //$$

Dirichlet's theorem shows that the structure of the unit group of $Q(\delta)$ depends on whether the field is real or complex.

COMPLEX $Q(\delta)$. The only units are the following roots of unity.

$d = -1$	$\pm 1, \pm i$
$d = -3$	$\pm 1, (\pm 1 \pm \sqrt{-3})/2$
$d \neq -1, -3$	± 1

REAL $Q(\delta)$ The units are of the form

$$\pm (\varepsilon(d))^k, \quad k \in Z$$

where $\varepsilon(d)$ is the unit satisfying $1 < \varepsilon(d)$, $\varepsilon(d)$ minimal. We shall refer to $\varepsilon(d)$ as *the* fundamental unit of $Z(\delta)$. $\varepsilon(d)$ is most efficiently calculated by the use of the simple continued fraction algorithm which is described in the next section.

Recall that the integers of $Z(\delta)$ do not necessarily factor uniquely but that unique factorization is "recovered" through the study of the ideals of $Z(\delta)$. We shall need the following terminology and results concerning the ideals of $Z(\delta)$ and their property of unique factorization.

We shall use $<\alpha_1,\alpha_2,\ldots,\alpha_k>$ to denote the $Z(\delta)$ ideal generated by $\alpha_1,\alpha_2,\ldots,\alpha_k \in Z(\delta)$. Recall that an ideal $I \subseteq Z(\delta)$ is said to be principal if and only if $\exists\ \alpha \in Z(\delta)$ such that $I = <\alpha>$. The conjugate ideal of an ideal $I \subseteq Z(\delta)$ will be denoted by I', that is

$$I' = \{\alpha' : \alpha \in I\}$$

Recall that if $I = I'$ then I is said to be self conjugate. The norm of an ideal $I \subseteq Z(\delta)$ will be denoted by $N(I)$. (Note that $N(I)$ is a positive rational integer and that

$$II' = <N(I)>$$

Thus II' is not only principal but is also the ideal generated by $N(I)$. Note also that $N(I) = N(I')$ and $N(<\alpha>) = |N(\alpha)|$.

Of course any ideal of $Z(\delta)$ factors uniquely as a product of prime ideals. For ease of reference we now note the prime ideals of $Z(\delta)$.

THEOREM 1.3 (Cohn [1962,pp142-145])

Let (Δ/p) be the Kronecker symbol. Then the prime ideals of $Z(\delta)$

are precisely those ideals of the following three types.

(i) $P = \langle p \rangle$ where p is a rational prime for which $(\Delta/p) = -1$

(ii) P, P' (distinct) with $PP' = \langle p \rangle$ where p is a rational prime for which $(\Delta/p) = 1$

(iii) P with $P^2 = \langle p \rangle$, $P = P'$ where p is a rational prime for which $(\Delta/p) = 0$, that is $p|\Delta$

(Note that type (i) ideals are principal while type (ii), (iii) ideals may or may not be principal, and type (i), (iii) ideals are self-conjugate while type (ii) ideals are not self-conjugate).

A more precise description of the prime ideals of $Z(\delta)$ is given by the following decompositions of the rational prime ideals. The factors in these decompositions give all the prime ideals of $Z(\delta)$.

$$\langle 2 \rangle = \begin{cases} \langle 2,\omega \rangle \langle 2,\omega' \rangle & \text{if } d \equiv 1 \pmod 8 \quad ((\Delta/2) = 1) \\ \langle 2 \rangle & \text{if } d \equiv 5 \pmod 8 \quad ((\Delta/2) = -1) \\ \langle 2,\delta \rangle^2 & \text{if } d \equiv 2 \pmod 4 \quad (2|\Delta) \\ \langle 2,1+\delta \rangle^2 & \text{if } d \equiv 3 \pmod 4 \quad (2|\Delta) \end{cases}$$

and for $p > 2$

$$\langle p \rangle = \begin{cases} \langle p \rangle & \text{if } (\Delta/p) = -1 \\ \langle p,a+\delta \rangle \langle p,a-\delta \rangle & \text{if } (\Delta/p) = 1 \quad (a^2 \equiv d \pmod p) \\ \langle p,\delta \rangle^2 & \text{if } p|\Delta \end{cases}$$ //

The study of ideals and unique factorization also leads to the concept of the class number of the field $Q(\delta)$ which we will denote by $h(d)$. Of course $Z(\delta)$ is a unique factorization domain if and only if $h(d) = 1$.

In this thesis we will often need to represent an ideal as a Z module.

THEOREM 1.4 (Cohn [1962, chapters 4,7,8])

The ideal $I \subseteq Z(\delta)$ can be represented in Z module form as

$$I = Z[a,\alpha]$$

where the basis $\{a,\alpha\}$ satisfies

(i) a is the minimal positive rational integer in I

(ii) $\alpha = (e+f\delta)/c \in Z(\delta)$ with f the minimal positive coefficient of δ occurring in I and $0 \leq e < ac$.

This representation is unique and will be referred to as the standard representation of I. Furthermore $N(I) = af$ and $f|e$, $f|a$. //

Note that $<\alpha> = Z[\alpha,\omega\alpha]$ although this is not usually the standard representation of $<\alpha>$. In this thesis we will often need to reduce a non standard ideal representation

$$I = Z[\alpha_1,\alpha_2,\ldots,\alpha_k],\ \alpha_1,\alpha_2,\ldots,\alpha_k \in Z(\delta)$$

to the standard representation for I. We remind the reader of the technique used with the following example.

EXAMPLE 1.1

Let $\delta = \sqrt{2}$, $\alpha = 4 + 3\delta$, $\beta = 6 + 7\delta$. Using a Euclidean type algorithm which first reduces the δ coefficients we have

$$\begin{aligned} <\alpha,\beta> &= Z[\alpha,\alpha\omega,\beta,\beta\omega] \\ &= Z[4+3\delta,6+4\delta,6+7\delta,14+6\delta] \\ &= Z[4+3\delta,2+\delta,-2+\delta,6] \\ &= Z[-2,2+\delta,-4,6] \\ &= Z[-2,\delta,0,0] \\ &= Z[2,\delta] \end{aligned}$$

which is the standard representation of $<\alpha,\beta>$ as a Z module. //

We shall also have need of the following result concerning the basis of a standard representation of an ideal.

THEOREM 1.5

Let

$$I = <\alpha_1,\alpha_2,\ldots,\alpha_k>,\ \alpha_j = a_j + b_j\omega,\ \ a_j, b_j \in Z$$

be an ideal of $Z(\delta)$. Then in the standard representation of I described in theorem 1.4 we have that the δ coefficient of α is

$$f = (a_1,b_1,a_2,b_2,\ldots,a_k,b_k)$$

PROOF

Let $f_1 = (a_1,b_1,\ldots,a_k,b_k)$. Clearly $\forall\ \beta \in I\ \ f_1|\beta$. In particular $f_1|\alpha = (e+f\delta)/c$ and it follows that $f_1|f$. (If $c = 2$ then $(e+f\delta)/2 = (e-f)/2 + f\omega$). Since $\omega'\alpha_j = b_jN(\omega) + a_j\omega'$ it is not difficult to see that we have $e_1 + f_1\omega \in I$ for some $e_1 \in Z$. Consequently $f|f_1$ and the result now follows. //

We finish this section on quadratic fields with a paragraph which only applies to complex quadratic fields, that is the subcase $d < 0$. For this subcase we have $\alpha' = \bar{\alpha}$. Note that $Z(\delta)$ is a discrete subset of $\mathbb{C}$. In fact if $\alpha,\beta \in Z(\delta)$, $\alpha \neq \beta$ then $|\alpha-\beta| \geq 1$. We can therefore generalize the nearest integer function as follows. Given $x \in \mathbb{C}$ we define

$$\{x\} = \alpha \in Z(\delta) \text{ where } |x-\alpha| \leq |x-\beta|\ \ \forall\ \beta \in Z(\delta) \qquad (3)$$

(Ties to be decided by some arbitrary rule). Of course the nearest integer function now depends on the complex quadratic field under consideration. Note that if $\alpha \in Z(\delta)\backslash Z$ then $|\alpha-\mathrm{Re}(\alpha)| \geq \sqrt{3}/2 > 1/2$ (the minimum value occurs when $d = -3$, $\alpha = (a\pm\delta)/2$). Consequently (1)

and (3) agree for $x \in \mathbb{R}$. We also define

$$Z(\delta)^+ = \{\alpha \in Z(\delta) : \alpha \neq 0, \arg(\alpha) \in (-\pi/2, \pi/2]\}$$

Note that for $\alpha \in Z(\delta)$ we have precisely one of either $\alpha = 0$, or $\alpha \in Z(\delta)^+$, or $-\alpha \in Z(\delta)^+$. Finally note that $h(d) = 1$ for

$$d = -1, -2, -3, -7, -11, -19, -43, -67, -163$$

and that $Q(\delta)$ is Euclidean for

$$d = -1, -2, -3, -7, -11.$$

SECTION THREE

SIMPLE CONTINUED FRACTIONS AND REAL QUADRATIC UNITS

The main purpose of this section is to briefly review material concerning the simple continued fraction algorithm with particular reference to its application to the problem of determining the fundamental unit of a real quadratic field. The ideas and results covered in this section plus the following section are central to much of the work in this thesis in that we will use them as a guide in our development of an algorithm for the calculation of units of quartic fields which are quadratic extensions of complex quadratic fields.

Detailed development of the theory reviewed below can be found in texts such as Hardy and Wright [1979], Chrystal [1959], and LeVeque [1977].

Throughout this section we assume that $d > 0$. Thus $\delta, \omega \in \mathbb{R}$ and the field $Q(\delta)$ is real.

A simple continued fraction will be denoted by

$$(a_1, a_2, \ldots)$$

The partial quotients $a_1, a_2, \ldots$ satisfy $a_1 \in Z$, $a_2, \ldots \in Z^+$. The convergents of a simple continued fraction will be denoted by

$$p_k/q_k, \quad k = -1, 0, 1, 2, \ldots$$

where

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

$$p_k = a_k p_{k-1} + p_{k-2}, \; q_k = a_k q_{k-1} + q_{k-2}, \; k \in Z^+$$

Successive convergents satisfy

$$p_k q_{k-1} - q_k p_{k-1} = (-1)^k \tag{4}$$

Of course every simple continued fraction represents a real number and conversely every real number can be represented as a simple continued fraction which can be obtained as follows.

ALGORITHM 1.1

Let $x \in \mathbb{R}$. Then the partial quotients of the simple continued fraction expansion of x are calculated as follows.

1 Set $x_1 = x$, $k = 1$

2 Set $a_k = [x_k]$

3 If $a_k = x_k$ then stop

4 Set $x_{k+1} = (x_k - a_k)^{-1}$

5 Increment k by 1, go to 2 //

Note that in practice the number of partial quotients which can be accurately calculated using algorithm 1.1 depends on the initial accuracy of the approximation to x. Consequently it is often necessary to use multiprecision arithmetic if we wish to calculate more than a dozen or so partial quotients.

The simple continued fraction expansion of x is essentially unique and we have

x has a finite simple continued fraction expansion

if and only if (5)

$x \in Q$

Thus if $x \in Q$ we have

$$x = (a_1, a_2, \ldots, a_n) \quad \text{and} \quad x = p_n/q_n$$

Note that if $p, q \in Z$, $(p,q) = 1$ then from (4) we see that a solution of $pa - qb = 1$ with $a, b \in Z$ is given by $a = (-1)^n q_{n-1}$, $b = (-1)^n p_{n-1}$ where $p/q = (a_1, a_2, \ldots, a_n)$.

The complete quotients x_k, $k = 1,2,\ldots$ in algorithm 1.1 satisfy

$$x_k = -(p_{k-2} - q_{k-2}x)/(p_{k-1} - q_{k-1}x) \tag{6}$$

The convergents of the simple continued fraction expansion of x form a sequence of increasingly better approximations to x. In fact we have

$$|x - p_k/q_k| \leq 1/q_k q_{k+1} \leq 1/q_k^2 \tag{7}$$

The first equality is only possible if $x = p_{k+1}/q_{k+1}$ and the second equality is only possible if $q_k = q_{k+1} = k = 1$. As a partial converse of (7) we have that if $x \in R$, $p, q \in Z$ then

$$|p/q - x| < 1/2q^2$$

implies (8)

p/q is a convergent of the simple continued fraction expansion of x.

More important than (7) as far as this thesis is concerned is the idea

of a best approximation. For $p \in Z$, $q \in Z^+$, $x - [x] \neq 1/2$ we say

p/q is a best approximation to x

if and only if (9)

$\forall\ a,b \in Z,\ 0 < b \leq q,\ a/b \neq p/q$ we have $|p-qx| < |a-bx|$

THEOREM 1.6 (LeVeque [1977, sections 9.2, 9.3])

If $x - [x] \in [0,1/2)$ (resp. $x - [x] \in (1/2,1)$) then the convergents p_k/q_k, $k = 1,2,\ldots$ (resp. $k = 2,3,\ldots$) of the simple continued fraction expansion of x are precisely the best approximations to x. (If $x \in Q$ then assume the expansion is the shorter of the two possibilities). //

A periodic simple continued fraction will be denoted by

$$(a_1,a_2,\ldots,a_{m-1},\ \overline{a_m,\ldots,a_{m+r-1}}) \tag{10}$$

We assume that $m,r \in Z^+$ are minimal. The partial quotients $a_1,\ldots,a_{m-1}$ form the preperiod and the partial quotients $a_m,\ldots,a_{m+r-1}$ form the period. If $m = 1$ then the simple continued fraction is said to be purely periodic. Note that for an expansion of form (10) we have

$$x_{k+r} = x_k,\ \forall\ k \geq m$$

Consequently it is easily deduced using (6) that if x has an expansion of form (10) then x is a real quadratic surd. The simple continued fraction expansion of any real quadratic surd is most easily calculated using a modified form of algorithm 1.1.

ALGORITHM 1.2 (Chrystal [1959, chapter 33])

Let $x = (a+b\delta)/e$, $a,e \in Z$, $b \in Z^+$ with $a^2 \equiv b^2 d \pmod{e}$. (If necessary this can be arranged by replacing a, b, e with $a|e|, b|e|, e|e|$). Then the simple continued fraction expansion of x can be

calculated as follows.

1 Set $P_1 = a$, $Q_1 = e$, $k = 1$

2 Set $a_k = [(P_k+b\delta)/Q_k]$

3 Set $P_{k+1} = a_k Q_k - P_k$

4 Set $Q_{k+1} = (b^2d-P_k^2)/Q_k$

5 Increment k by 1, go to 2 //

We have

$$x_k = (P_k+b\delta)/Q_k, \quad k = 1,2,... \tag{11}$$

and P_k, Q_k are small rational integers which eventually satisfy

$$0 < P_k, Q_k < 2b\delta$$

It follows that the x_k can only assume finitely many distinct values and so the expansion of any real quadratic surd is periodic with the end of the first period corresponding to the first occurrence of $x_{m+r} = x_m$, that is $P_{m+r} = P_m$, $Q_{m+r} = Q_m$. We therefore have the result

x has a periodic simple continued fraction expansion

if and only if

x is a real quadratic surd.

Note that it is easily deduced from (6) and (11) that

$$Q_k = (-1)^{k-1}Q_1N(p_{k-1}-q_{k-1}x), \quad k = 1,2,... \tag{12}$$

Since algorithm 1.2 basically only involves integer arithmetic ([bδ] is a good enough approximation to bδ in step 2) the advantage this algorithm has over algorithm 1.1 from a computational point of view is obvious.

We now review the connection between real quadratic units and simple continued fractions. Using (8) it is relatively easy to show that if $\varepsilon(d) = a - b\omega'$ then for $d \neq 5$ a/b is a convergent of the simple continued fraction expansion of ω. In view of (12) we have for $d \neq 5$ that $\varepsilon(d) = p_{j-1} - q_{j-1}\omega'$ where $j > 1$ is the minimal integer for which $Q_j = Q_1$. A study of the coefficients obtained in the expansion of ω using algorithm 1.2 enables us to be more specific about this value of j.

THEOREM 1.7 (Chrystal [1959, chapter 33])

For $d \neq 5$ the sequences P_k, Q_k, a_k obtained when ω is expanded by algorithm 1.2 exhibit the following symmetries.

k	1	2	3	4 ... r-1	r	r+1	r+2	...
P_k	P_1	P_2	P_3	$P_4 \cdots P_4$	P_3	P_2	P_2	...
Q_k	Q_1	Q_2	Q_3	$Q_4 \cdots Q_3$	Q_2	Q_1	Q_2	...
a_k	a_1	a_2	a_3	$a_4 \cdots a_3$	a_2	a_{r+1}	a_2	...

We have

$$\omega = (a_1, \overline{a_2, a_3, \ldots, a_3, a_2, a_{r+1}}), \quad a_{r+1} = \begin{cases} 2a_1 & \text{if } \omega = \delta \\ 2a_1 - 1 & \text{otherwise} \end{cases}$$

$$0 \leq P_k < \delta, \quad 1 \leq Q_k < 2\delta, \quad 1 \leq a_k < 2\delta$$

Furthermore

$$Q_j = Q_1 \text{ if and only if } \quad j \equiv 1 \pmod{r}$$

and so

$$\varepsilon(d) = p_r - q_r\omega' \qquad //$$

The amount of work involved in calculating $\varepsilon(d)$ can be approximately halved if we take advantage of the symmetric nature of the P_k, Q_k in theorem 1.7.

THEOREM 1.8

For $d \neq 5$ the midpoint of the expansion of ω can be recognized by the first occurrence of one of the following conditions.

(a) $Q_{k+1} = Q_k, a_{k+1} = a_k$, in which case $r = 2k - 1$,

$$P_{k+2} = P_k,\ Q_{k+2} = Q_{k-1},\ a_{k+2} = a_{k-1}$$

$$P_{k+3} = P_{k-1},\ Q_{k+3} = Q_{k-2},\ a_{k+3} = a_{k-2} \text{ etc,}$$

and

$$\varepsilon(d) = (p_{k-1}-q_{k-1}\omega')(p_k-q_k\omega')/|N(p_{k-1}-q_{k-1}\omega')| \tag{13}$$

(b) $P_{k+1} = P_k$ in which case $r = 2k - 2$,

$$Q_{k+1} = Q_{k-1},\ a_{k+1} = a_{k-1}$$

$$P_{k+2} = P_{k-1},\ Q_{k+2} = Q_{k-2},\ a_{k+2} = a_{k-2} \text{ etc,}$$

and

$$\varepsilon(d) = (p_{k-1}-q_{k-1}\omega')^2/|N(p_{k-1}-q_{k-1}\omega')| \tag{14}$$

PROOF

Chrystal [1959, chapter 33] gives the results concerning the P_k, Q_k, a_k. To prove (13) and (14) we use (6) to show that

$$\varepsilon(d) = p_r - q_r\omega' = x_2 x_3 \cdots x_{r+1}$$

and then use the symmetries exhibited by the P_k, Q_k to obtain the expressions given for $\varepsilon(d)$. //

At this point we note that if $d \equiv 1 \pmod 4$, $d \neq 5$ then $\varepsilon(d)$ can also be calculated using the simple continued fraction expansion of δ.

THEOREM 1.9

Let $d \equiv 1 \pmod 4$. Then theorem 1.7 and the symmetries of theorem

1.8 also apply to the expansion of δ. $\varepsilon = p_r + q_r\delta$ is the fundamental unit of $Z[1,\delta] \subsetneq Z(\delta)$. For $d > 5$ we have

(a) if $d \equiv 1 \pmod 8$ then $\varepsilon = \varepsilon(d)$

(b) if $d \equiv 5 \pmod 8$ then $\varepsilon = \varepsilon(d)$ or $(\varepsilon(d))^3$. If $\varepsilon = (\varepsilon(d))^3$ then $\exists\ j, k \in Z^+$, $1 < j < k \leq r$ such that $Q_j = Q_k = 4$ and

$$\varepsilon(d) = (p_{j-1}+q_{j-1}\delta)/2,\ (\varepsilon(d))^2 = (p_{k-1}+q_{k-1}\delta)/2 \qquad //$$

We finish this section by noting a variation on algorithm 1.1 which is often referred to as the nearest integer continued fraction algorithm.

ALGORITHM 1.3

Let $x \in \mathbb{R}$. Then the partial quotients of the nearest integer continued fraction expansion of x are calculated as follows.

1 Set $x_1 = x$, $k = 1$

2 Set $a_k = \{x_k\}$

3 If $a_k = x_k$ then stop

4 Set $x_{k+1} = (x_k - a_k)^{-1}$

5 Increment k by 1, go to 2 //

The properties of the nearest integer continued fraction expansion of x are similar to those of the simple continued fraction expansion of x. However the nearest integer continued fraction algorithm is not guaranteed to obtain all best approximations to x. The nearest integer continued fraction expansion of x is periodic if and only if x is a real quadratic surd and the obvious modification of algorithm 1.2 applies. We can also use the nearest integer continued fraction expansion of ω to calculate $\varepsilon(d)$ and this is generally more efficient than using the simple continued fraction expansion of x. (See Williams and Buhr [1979]). However the work in chapter two is more

directly related to the better known ideas of the simple continued fraction algorithm and best approximations. It is for this reason that we have not featured the nearest integer continued fraction algorithm more prominently.

SECTION FOUR

RELATIVE MINIMA OF Z MODULES

In this section we present an alternative approach to the material in the previous section which is based on the ideas of a best approximation given in (9). The ideas presented are basically due to Voronoi [1896] and they provide the most appropriate basis for the generalization which we will be considering in chapter two.

For the set $\mathbb{R} \times \mathbb{R}$ we can define arithmetic operations in the obvious componentwise fashion. For example if $A = (x,y)$, $B = (z,t) \in \mathbb{R} \times \mathbb{R}$ then $AB = (xz,yt)$ and $A/B = (x/z,y/t)$ (provided $zt \neq 0$). For $A = (x,y)$ we define

$$A^* = (y,x), \quad |A| = |x|$$

Clearly $(A/B)^* = A^*/B^*$, $|A^*| = |y|$, $|A/B| = |A|/|B|$.

For $x \in \mathbb{R}\setminus\{0\}$ let

$$M(x) = Z[(1,1),X], \quad X = (x,-x)$$

with module operations as defined in the previous paragraph. Note that for $A = a(1,1) + bX = (a+bx,a-bx) \in M(x)$ we have

$$A^* = a(1,1) - bX = (a-bx,a+bx), \quad |A^*| = |a-bx|$$

$$A + A^* = 2a(1,1), \quad A - A^* = 2b(x,-x)$$

A relative minimum of $M(x)$ is any $A \in M(x)\backslash\{(0,0)\}$ satisfying

$$\forall\ B \in M(x)\backslash\{(0,0)\}\ |B| < |A| \text{ implies } |B^*| > |A^*|$$

$$\text{and } |B| = |A| \text{ implies } |B^*| \geq |A^*|$$

The following results are easily proved.

THEOREM 1.10

(a) If $A = a(1,1) + bX \in M(x)$ is a relative minimum

then $(a,b) = 1$.

(b) $A_0 = \begin{cases} (1,1) & |x| \geq 1 \\ X & |x| < 1 \end{cases}$ is a relative minimum of $M(x)$.

(c) If A is a relative minimum of $M(x)$ then so are $-A, \pm A^*$. Furthermore if $|x| \neq 1$ then $A, -A$ are the only relative minima of magnitude $|A|$.

(d) If A is a relative minimum of $M(x)$ then either $|A| = |A_0|$, or $|A| > |A_0| > |A^*|$, or $|A| < |A_0| < |A^*|$. //

A relative minimum $A = (p+qx, p-qx)$ for which $q \in Z^+$ and $|A| > |A_0|$ will be called a positive relative minimum. In view of theorem 1.10 it is easily seen that we can restrict our attention to the positive relative minima of $M(x)$. $M(x)$ is a countable set and so we can label the positive relative minima as A_k, $k = 1,2,\ldots$ so that we have $|A_0| < |A_1| < |A_2| < \ldots$ For $k = 0,1,2,\ldots$ and $|A_k^*| \neq 0$ we define

$$R_k = A_{k+1}/A_k,\ M_k(x) = M(x)/A_k = \{A/A_k : A \in M(x)\}$$

THEOREM 1.11

(a) We can find $B_k \in M(x)$ (not unique) such that

$$M(x) = Z[A_k, B_k]$$

and so

$$M_k(x) = Z[(1,1),X_k],\ X_k = B_k/A_k$$

(b) R_k is the unique $R \in M_k(x)$ having negative first component if $k = 0$, $x \leq -1$ and positive first component otherwise which also satisfies $|R^*| < 1$, $|R|$ minimal.

(c) We can find $S_k \in M_k(x)$ such that

$$M_k(x) = Z[R_k,S_k]$$

and so

$$M_{k+1}(x) = M_k(x)/R_k = Z[(1,1),X_{k+1}],\ X_{k+1} = S_k/R_k \qquad //$$

This theorem suggests the following method for calculating positive relative minima.

ALGORITHM 1.4

Let $x \in \mathbb{R}\setminus\{0\}$. Then the positive relative minima of $M(x)$ can be calculated as follows.

1 If $|x| \geq 1$ then set $A_0 = (1,1)$, $X_0 = X$

else set $A_0 = X$, $X_0 = X^{-1}$

2 Set $k = 0$

3 Find $R_k \in M_k(x) = Z[(1,1),X_k]$

4 Set $A_{k+1} = R_k A_k$

5 If $|R_k^*| = 0$ stop

6 Find $S_k \in M_k(x)$ such that $M_k(x) = Z[R_k,S_k]$

7 Set $X_{k+1} = S_k/R_k$, that is $M_{k+1}(x) = Z[(1,1),X_{k+1}]$

8 Increment k by 1, go to 3 //

In fact algorithm 1.4 simplifies to what is essentially the simple continued fraction algorithm. (In the simplest case $x \geq 1$ we find

$R_k = a_{k+1}(1,1) + X_k$ and we can take $S_k = (1,1)$. It follows that $A_k = p_k(1,1) + q_k X$, $X_k = A_{k-1}/A_k$, $k = 1,2,\ldots$, and

$$X_k = \left(\frac{p_{k-1}+q_{k-1}x}{p_k+q_k x} , \frac{p_{k-1}-q_{k-1}x}{p_k-q_k x} \right)$$

$$= \left(\frac{p_{k-1}+q_{k-1}x}{p_k+q_k x} , -x_{k+1} \right) \qquad k = 0,1,2,\ldots$$

For $x < 1$ there are several minor differences). However we have given algorithm 1.4 in a semi-developed form because it is this form which is the basis for generalization in chapter two. The simplified form of algorithm 1.4 (that is the simple continued fraction algorithm) is much less suitable for this purpose and we shall not pursue the precise details of this simplification. However we finish this section by noting the precise connection between the positive relative minima of $M(x)$ and the convergents of the simple continued fraction expansion of x. Note that this theorem shows that the results of section three (after appropriate modification) apply to the positive relative minima of $M(x)$.

THEOREM 1.12

Let $x \in \mathbb{R}\backslash\{0\}$. Let p_k/q_k denote the k^{th} convergent of the simple continued fraction expansion of x. Then the positive relative minima of $M(x)$ are as follows. (If $x \in Q$ then assume the expansion is the shorter of the two possibilities).

(a) $x \geq 1$ $\quad A_k = p_k(1,1) + q_k X, \quad k = 1,2,\ldots$

(b) $0 < x < 1$ $\quad A_k = p_{k+1}(1,1) + q_{k+1} X, \quad k = 1,2,\ldots$

(c) $-\frac{1}{2} < x < 0$ $\quad A_k = p_{k+2}(1,1) + q_{k+2} X, \quad k = 1,2,\ldots$

(d) $x = -\frac{1}{2}$ $\quad A_1 = -(1,1) + 2X$

(e) $-1 \leq x < -\frac{1}{2}$ $\quad A_k = p_k(1,1) + q_k X \quad k = 1,2,\ldots$

(f) $x < -1$

(i) $[x] + \frac{1}{2} < x < [x] + 1$ $\quad A_k = p_{k+1}(1,1) + q_{k+1}X, \; k = 1,2,\ldots$

(ii) $x = [x] + \frac{1}{2}$ $\quad A_1 = (p_1+1)(1,1) + q_1X$

$\quad A_2 = p_2(1,1) + q_2X$

(iii) $[x] < x < [x] + \frac{1}{2}$ $\quad A_1 = (p_1+1)(1,1) + q_1X$

$\quad A_k = p_{k-1}(1,1) + q_{k-1}X \quad k = 2,3,\ldots$

(iv) $x = [x]$ $\quad A_1 = p_1(1,1) + q_1X$ //

CHAPTER TWO

RELATIVE MINIMA OF MODULES OVER A RING OF COMPLEX QUADRATIC INTEGERS

In chapter one section four we noted how the idea of relative minima of modules of the form $Z[(1,1),X]$, $X = (x,-x)$, $x \in \mathbb{R}\backslash\{0\}$ leads to what is essentially the simple continued fraction algorithm and hence (chapter one section three) to a method for calculating fundamental units of real quadratic fields. In this chapter we shall generalize this idea to modules of the form $Z(\delta)[(1,1),W]$ where $W = (w,-w)$, $w \in \mathbb{C}\backslash\{0\}$, and $Z(\delta)$ is a ring of complex quadratic integers. Our main motivation for considering this idea is the desire to develop an algorithm which can be used to calculate fundamental units of quadratic extensions of complex quadratic fields. However in chapter three we shall see that the work in this chapter is not entirely successful in this respect although fortunately the problems which arise do not prove to be insurmountable.

SECTION ONE

DEFINITIONS, NOTATIONS, AND BASIC THEOREMS

Throughout this chapter we assume that $Q(\delta)$ is a complex quadratic field (see chapter one section two).

We shall need the following operations on $\mathbb{C} \times \mathbb{C}$.

DEFINITION 2.1

Let $(u,v),(x,y) \in \mathbb{C} \times \mathbb{C}$, $z \in \mathbb{C}$. We define

(a) $(u,v) \pm (x,y) = (u \pm x, v \pm y)$

(b) $(u,v)(x,y) = (ux,vy)$

(c) $(u,v)/(x,y) = (u/x, v/y)$ provided $xy \neq 0$

(d) $z(u,v) = (zu,zv)$

(e) $(u,v)^* = (v,u)$

called the *reverse* of (u,v)

(f) $|(u,v)| = |u|$

called the *magnitude* of (u,v)

(g) $\overline{(u,v)} = (\bar{u},\bar{v})$

called the *complex conjugate* of (u,v) //

Note that the reverse function preserves operations a,b,c,d and the magnitude function preserves operations b,c,d. Furthermore the magnitude function satisfies the triangle inequality and if $A \in \mathbb{C} \times \mathbb{C}$ then $|AA^*| \neq 0$ if and only if neither component of A is zero. Finally note that $(A^*)^* = A$ and that AA^* has identical components.

In this chapter we will be primarily interested in the properties of the following $Z(\delta)$ module which is a subset of $\mathbb{C} \times \mathbb{C}$.

DEFINITION 2.2

Let $w \in \mathbb{C} \setminus \{0\}$. We define $M(w)$ to be the module

$$M(w) = Z(\delta)[(1,1),W], \quad W = (w,-w)$$

where $Z(\delta)$ is a ring of complex quadratic integers. The module operations are given in definition 2.1. For $A \in M(w)$ we have

$$A = \alpha(1,1) + \beta W = (\alpha+\beta w, \alpha-\beta w), \quad \alpha,\beta \in Z(\delta)$$

The $Z(\delta)$ integers α,β will be referred to as the coefficients of A. //

Clearly $A \in M(w)$ implies $A^* \in M(w)$. Since $(1,1)$, W are linearly independent over $\mathbb{C}$ we see that the coefficients of $A \in M(w)$ are unique. It is also easily seen that $\{A,B\}$ is a basis of $M(w)$ if and only if $A = \alpha(1,1) + \beta W$, $B = \kappa(1,1) + \lambda W \in M(w)$ with $\alpha\lambda - \beta\kappa = \xi$ a root of unity of $Z(\delta)$. Note that $\{A,\xi^{-1}B\}$ is also a basis whenever $\{A,B\}$ is a basis and so we can in general assume that $\xi = 1$.

The following theorem gives a number of fairly obvious results which will be used frequently in this chapter.

THEOREM 2.1

(a) Let $A = \alpha(1,1) + \beta W \in M(w)$. Then

$$|A| = |\alpha+\beta w|,\ A + A^* = 2\alpha(1,1),\ A-A^* = 2\beta(w,-w)$$

$$2|\alpha| - |A^*| \leqq |A| \leqq 2|\alpha| + |A^*|$$

$$2|\beta w| - |A^*| \leqq |A| \leqq 2|\beta w| + |A^*|$$

(b) Let $r,s \in \mathbb{R}^+$. Then there are only finitely many $A \in M(w)$ satisfying $|A| < r$, $|A^*| < s$.

PROOF

Only part (b) requires some explanation. Suppose $A = \alpha(1,1) + \beta W$ satisfies $|A| < r, |A^*| < s$. Then from part (a) we have

$$2|\alpha|, 2|\beta w| \leq |A| + |A^*| < r + s$$

The result is now clear since there are only finitely many $\alpha,\beta \in Z(\delta)$ satisfying these conditions. //

In developing results concerning $M(w)$ we shall need to consider a larger class of $Z(\delta)$ modules which are derived from $M(w)$.

DEFINITION 2.3

Let $X \in \mathbb{C}X\mathbb{C}$ and suppose $|XX^*| \neq 0$. Then we define

$$M(w)/X = Z(\delta)[(1,1)/X, W/X]$$

$$= \{B/X : B \in M(w)\}$$

The module operations are given in definition 2.1. Furthermore if $Y \in \mathbb{C}X\mathbb{C}$, $|YY^*| \neq 0$ then we define

$$(M(w)/X)/Y = M(w)/(XY) \qquad //$$

We will mainly be interested in the cases where $X \in M(w)$ and $Y \in M(w)/X$. Note that $\{R,S\}$ is a basis of $M(w)/X$ if and only if $\{A = RX, B = SX\}$ is a basis of $M(w)$. However note that $(A/X)^* = A^*/X^*$ is not necessarily an element of $M(w)/X$.

We are now able to define a relative minimum of $M(w)/A$, $A \in M(w)$.

DEFINITION 2.4

Let $A \in M(w)$, $|AA^*| \neq 0$. A *relative minimum* of $M(w)/A$ is any $R \in (M(w)/A)\backslash\{(0,0)\}$ satisfying

$$\forall\, S \in (M(w)/A)\backslash\{(0,0)\} \quad |S| < |R| \text{ implies } |S^*| > |R^*|$$

$$\text{and } |S| = |R| \text{ implies } |S^*| \geq |R^*| \qquad //$$

Of course our main interest lies with the relative minima of $M(w)$. (That is $A = (1,1)$ in definition 2.4). However when calculating relative minima of $M(w)$ we shall find it convenient to work with relative minima of certain $M(w)/A$. The connection between relative minima of $M(w)$ and relative minima of $M(w)/A$ is simple and straightforward.

THEOREM 2.2

Let $B \in M(w)$, $|BB^*| \neq 0$. Then

A is a relative minimum of $M(w)$

if and only if

A/B is a relative minimum of $M(w)/B$

PROOF

The result is clear once we have noted the one to one correspondence

$$A \in M(w) \longleftrightarrow A/B \in M(w)/B$$

and relationships such as

$$|C/B| < |A/B| \text{ if and only if } |C| < |A|, \text{ etc.} \qquad //$$

The importance of this result with respect to the calculation of relative minima of M(w) will become clear in section three. However this result is also of more immediate use in that it simplifies some of the following proofs.

The next theorem shows that the relative minima of M(w), M(-w), $M(w^{-1})$, and $M(\bar{w})$ are closely related.

THEOREM 2.3

(a) M(w) and M(-w) have identical relative minima

(b) A is a relative minimum of M(w)

if and only if

A/W is a relative minimum of $M(w^{-1})$

(c) A is a relative minimum of M(w)

if and only if

$\bar{A}$ is a relative minimum of $M(\bar{w})$

PROOF

(a) The result is clear once we have noted that

$$\alpha(1,1) + \beta(w,-w) = \alpha(1,1) + (-\beta)(-w,w)$$

since this shows that M(w) = M(-w).

(b) We have

$$\begin{aligned} M(w)/W &= Z(\delta)[(1,1)/W, W/W] \\ &= Z(\delta)[(w^{-1},-w^{-1}),(1,1)] \\ &= M(w^{-1}) \end{aligned}$$

The result now follows from theorem 2.2. (Take $B = W$).

(c) The result is clear once we have noted that $A = \alpha(1,1) + \beta W \in M(w)$ implies $\bar{A} = \bar{\alpha}(1,1) + \bar{\beta}\bar{W} \in M(\bar{w})$ and that $|A| = |\bar{A}|$. //

So far we have developed some of the relationships that exist between the relative minima of distinct modules. We now turn our attention to relationships between and properties of the relative minima of a given module. The first result is a fairly trivial consequence of definition 2.4. However we shall refer to it on a number of occasions.

THEOREM 2.4

If R,S are relative minima of M(w)/A then

$$|R| > |S| \text{ implies } |R^*| < |S^*|$$ //

THEOREM 2.5

A/B is a relative minimum of M(w)/B

if and only if

A*/B is a relative minimum of M(w)/B

PROOF

In view of theorem 2.2 we need only consider the case $B = (1,1)$. Furthermore since $(A^*)^* = A$ we need only prove the forward implication.

Suppose A is a relative minimum of M(w). If A* is not a relative minimum of M(w) then we can find $C \in M(w)\backslash\{(0,0)\}$ satisfying

$$\text{either} \quad |C| < |A^*| \text{ and } |C^*| \leq |A|$$
$$\text{or} \quad |C| = |A^*| \text{ and } |C^*| < |A|$$

However $D = C^*$ therefore satisfies

$$\text{either} \quad |D| < |A| \text{ and } |D^*| \leq |A^*|$$
$$\text{or} \quad |D| = |A| \text{ and } |D^*| < |A^*|$$

which contradicts the fact that A is a relative minimum. //

Our next objective is to establish an ordering and labelling of the relative minima of M(w). We begin by noting that M(w) is a countable set and so there are at most countably many relative minima of M(w). The next theorem gives us a base point for our labelling system.

THEOREM 2.6

(a) (1,1) is a relative minimum of M(w)

if and only if

$$|w| \geq 1$$

(b) W is a relative minimum of M(w)

if and only if

$$|w| \leq 1$$

PROOF

(a) Suppose (1,1) is a relative minimum of M(w). If $|w| < 1$ then we have $|W| = |w| < 1 = |(1,1)|$ and $|W^*| = |-w| < 1 = |(1,1)^*|$. This contradicts the fact that (1,1) is a relative minimum.

Now suppose that $|w| \geq 1$. If (1,1) is not a relative minimum then we can find $A = \alpha(1,1) + \beta W \in M(w) \setminus \{(0,0)\}$ such that either $|A| < 1$ and $|A^*| \leq 1$ or $|A| \leq 1$ and $|A^*| < 1$. In either case $2|\alpha| = |A+A^*| < 2$ implies $|\alpha| < 1$, and $2|\beta w| = |A-A^*| < 2$ implies $|\beta| < 1$. Consequently $\alpha = \beta = 0$ and so $A = (0,0)$ which is a contradiction.

(b) This case is proved in a manner similar to part (a). //

DEFINITION 2.5

We define

$$A_0 = \begin{cases} (1,1) & \text{if } |w| \geq 1 \\ W & \text{if } |w| < 1 \end{cases}$$ //

Before extending this labelling to other relative minima of M(w)

we must consider several points. Note that if A is a relative minimum then so is UA where U = (ξ,ξ), ξ a $Z(\delta)$ root of unity. ($|UA| = |A|, |(UA)^*| = |A^*|$). Furthermore it is possible to have relative minima $A,B \in M(W)$ with $|A| = |B|$, $|A^*| = |B^*|$ yet A,B not related by a root of unity factor. For example when $|w| = 1$, w not a root of unity we find that (1,1), W are two such relative minima. (See A_3, $A_3^{(1)}$ in example 2.7b for a less trivial example).

DEFINITION 2.6

Let R,S be relative minima of M(w)/A satisfying

$$|R| = |S|, \quad |R^*| = |S^*|$$

We shall call R,S *equivalent relative minima* and write

$$R \sim S \qquad //$$

Clearly $R \sim S$ if and only if $R^* \sim S^*$. Of course this equivalence relation divides the relative minima of M(w) into equivalence classes. This fact is acknowledged in the following definition which also gives the required ordering and labelling of the relative minima of M(w).

DEFINITION 2.7

The sets of equivalent relative minima of M(w) will be denoted by E_k, $k \in Z$, as defined below. A_k for $k \neq 0$ will denote any element of E_k.

E_0 is the set of all relative minima equivalent to A_0.

E_k for $k > 0$ is the set of all equivalent relative minima

$$\text{satisfying } |A_k| > |A_{k-1}|, \ |A_k| \text{ minimal.}$$

E_k for $k < 0$ is the set of all equivalent relative minima

$$\text{satisfying } |A_k| < |A_{k+1}|, \ |A_k| \text{ maximal.}$$

(If E_k is empty for some $k > 0$ then we define $E_{k+1}, E_{k+2}, \ldots$ to be the empty set. Similarly for $k < 0$). Where we wish to denote more than one element of E_k we will use the notation $A_k, A_k^{(1)}, A_k^{(2)}, \ldots$ The coefficients of $A_k, A_k^{(j)}$ will be denoted by $\alpha_k, \beta_k, \alpha_k^{(j)}, \beta_k^{(j)}$, that is

$$A_k = \alpha_k(1,1) + \beta_k W, \quad A_k^{(j)} = \alpha_k^{(j)}(1,1) + \beta_k^{(j)} W$$

Any sequence of relative minima $\ldots, A_k, A_{k+1}, A_{k+2}, \ldots$ (finite or infinite) will be called a *chain* of relative minima. A chain which contains an element from every non-empty E_k will be called a *complete chain* and a chain which contains A_0 and an element from every non-empty E_k for $k > 0$ will be called a *half chain*. //

It is fairly obvious that each E_k is a finite set (see theorem 2.1b) and with a little effort we can obtain a more precise upper bound for the number of elements in E_k. Divide the complex plane into the 6 regions

$$R_j = \{x \in \mathbb{C} : \arg(x) \in [(j-1)\pi/3, j\pi/3)\}, \quad j = 1,2,\ldots,6$$

Now if E_k has 37 or more elements then we can find $A_k, A_k^{(1)}$ such that $\alpha_k + \beta_k w, \alpha_k^{(1)} + \beta_k^{(1)} w \in R_j$ and $\alpha_k - \beta_k w, \alpha_k^{(1)} - \beta_k^{(1)} w \in R_m$. However it then follows that $A = A_k - A_k^{(1)}$ satisfies $|A| < |A_k|$, $|A^*| < |A_k^*|$ which contradicts the fact that A_k is a relative minimum. Thus E_k has at most 36 elements. However we generally find that the number of elements in E_k is the same as the number of roots of unity in $Z(\delta)$.

Note that the elements of a complete chain of relative minima satisfy

$$\ldots < |A_{-1}| < |A_0| < |A_1| < \ldots \tag{1}$$

Furthermore if $\{A_k\}$ is a complete chain of relative minima of $M(w)$ then theorem 2.2 shows that $\{A_k/A\}$ is a complete chain of relative minima of

M(w)/A. (Although definition 2.7 only applies to M(w) the idea of a complete chain extends easily to M(w)/A).

The next theorem shows that we can restrict our attention to the relative minima of M(w) having magnitude greater than $|A_0|$.

THEOREM 2.7

The sets of equivalent relative minima of M(w) satisfy

$$E_{-k} = \{A^* : A \in E_k\}$$

That is $A_k^* \sim A_{-k}$. Consequently the half chain $A_0, A_1, A_2, \ldots$ can be extended into a complete chain by taking A_{-k} to be A_k^* for $k = 1,2,\ldots$

PROOF

From definition 2.7 and theorem 2.4 we have (1) and

$$\ldots < |A_1^*| < |A_0^*| < |A_{-1}^*| < \ldots$$

Furthermore $\{A_k^*\}$ is a complete chain whenever $\{A_k\}$ is a complete chain. Since $|A_0| = |A_0^*|$ the theorem is now clear. //

In section three of this chapter we shall consider the problem of calculating a half chain of relative minima of M(w). Note that when calculating relative minima of M(w) we will initially only be interested in calculating one arbitrary representative of each E_k, $k > 0$.

The next theorem will be of importance when we develop an algorithm for calculating relative minima in section three. However we present it at this point so that we can use it in its simplest form (A = (1,1)) in example 2.1 which follows the theorem.

THEOREM 2.8

The relative minimum $A_{k+1}/A \in M(w)/A$ satisfies

(a) $|(A_{k+1}/A)^*| < |(A_k/A)^*|$

(b) $\forall\ B/A \in (M(w)/A)\backslash\{(0,0)\}$ satisfying $|(B/A)^*| < |(A_k/A)^*|$

we have either $|B/A| > |A_{k+1}/A|$

or $|B/A| = |A_{k+1}/A|$ and $|(B/A)^*| \geq |(A_{k+1}/A)^*|$

PROOF

In view of theorem 2.2 and the fact that $|A|, |A^*|$ are simply scale factors in the above inequalities we need only consider the case $A = (1,1)$. Furthermore condition (a) is just theorem 2.4.

To see that condition (b) must also hold we suppose to the contrary that we have $B \in M(w)\backslash\{(0,0)\}$ satisfying $|B^*| < |A_k^*|$ with $|B| \leq |A_{k+1}|$ and either $|B| \neq |A_{k+1}|$ or $|B^*| < |A_{k+1}^*|$. Now if $|B| = |A_{k+1}|$ then we must have $|B^*| < |A_{k+1}^*|$ which contradicts the fact that A_{k+1} is a relative minimum. Therefore we must have $|B| < |A_{k+1}|$. Now choose $C \in M(w)\backslash\{(0,0)\}$ satisfying $|C| \leq |B|$, $|C^*| \leq |B^*|$, $|CC^*|$ minimal. (Theorem 2.1b guarantees that the choice is from a finite set). It is easily checked that C is a relative minimum of M(w). Note that $|C^*| \leq |B^*| < |A_k^*|$ and so theorem 2.4 implies $|C| > |A_k|$. Thus C is a relative minimum of M(w) satisfying $|A_k| < |C| < |A_{k+1}|$ which is a contradiction. Therefore (b) must indeed hold. //

We now illustrate some of the ideas presented so far in this chapter. This example also suggests some of the ideas which we will develop in later sections.

EXAMPLE 2.1

Let $\delta = \sqrt{-10}$ and $w = \sqrt{1+\delta}$. We shall calculate a chain $A_0, A_1, \ldots, A_7$ in M(w) and note some of the more interesting points about these relative minima.

Of course $A_0 = (1,1)$ by definition since $|w| \geq 1$. A simple though very inefficient method for finding the remaining relative minima is to

perform an exhaustive search in the manner outlined in the following paragraph.

Suppose we know $A_k (k \geq 0, |A_k^*| > 0)$ and we wish to find A_{k+1}. From theorem 2.8 we see that we can take A_{k+1} to be any $A \in M(w)\backslash\{(0,0)\}$ which satisfies

(a) $|A^*| < |A_k^*|$, $|A|$ minimal

(b) $|B| = |A|$ implies $|B^*| \geq |A^*|$

We therefore search through the $\beta \in Z(\delta)^+$ (see the final paragraph of chapter one section two) in order of increasing magnitude until we find a β for which there exists an $\alpha \in Z(\delta)$ satisfying $|\alpha-\beta w| < |A_k^*|$. (Where there is more than one choice for α we choose the one which minimizes $|\alpha+\beta w|$). Now $A = \alpha(1,1) + \beta W$ may or may not be the required A_{k+1}. However $|A^*| < |A_k^*|$ and so theorem 2.8 gives $|A_{k+1}| \leq |A|$. Therefore by theorem 2.1a

$$2|\beta_{k+1}w| - |A_{k+1}^*| \leq |A_{k+1}| \leq |A| \leq 2|\beta w| + |A^*|$$

Since $|A_{k+1}^*|, |A^*| < |A_k^*|$ this implies

$$|\beta_{k+1}| < |\beta| + |A_k^*|/|w|$$

Therefore by checking the remaining possibilities indicated by this bound we will be able to determine A_{k+1}. Note that $|A_k^*| \leq |A_0| \leq |w|$ and so $|A_k^*|/|w| \leq 1$.

The search outlined produces the results given in table 2.1. (The calculations were of course done by a computer. See example 2.3). Note that the number of $\beta \in Z(\delta)^+$ satisfying $|\beta| < r$ is proportional to r^2. Consequently the amount of work required to find A_k quickly becomes unmanageable as k grows. Indeed finding the relative minima in table 2.1 requires the testing of approximately 10^6 possibilities.

TABLE 2.1 Relative minima of $M(\sqrt{1+\sqrt{-10}})$

k	α_k	β_k	$\lvert A_k\rvert^{\dagger}$	$\lvert A_k^*\rvert\lvert\beta_k\rvert^{\dagger}$
0	1	0	1.0	0
1	$4+\delta$	3	10.6	1.2
2	$9+2\delta$	6	21.9	1.4
3	$2+5\delta$	$6+2\delta$	31.8	.6
4	$37-14\delta$	$2-10\delta$	115.4	1.3
5	$104-3\delta$	$43-12\delta$	208.9	1.2
6	$245-20\delta$	$88-34\delta$	506.1	.3
7	$2464+327\delta$	$1427-108\delta$	5344.4	1.2

† Rounded to 1D

In section three we will develop a more efficient algorithm for calculating relative minima. This algorithm will still use the above search technique but on a fairly minor scale. In fact finding each A_k will involve a search which is roughly equivalent to the search required to find A_1 in the above example.

A second point to note is the fact that $|A_k^*||\beta_k|$ is approximately constant or equivalently $|A_k^*|$ is approximately proportional to $1/|\beta_k|$. In section two of this chapter we shall investigate this relationship in greater detail.

The most important observations about these relative minima come when we note that $\alpha_k + \beta_k w$, $\alpha_k - \beta_k w$ are conjugate algebraic integers in the quartic field $Q(w)$ which is clearly a quadratic extension of $Q(\delta)$. For $\alpha + \beta w \in Q(w)$ $(\alpha,\beta \in Q(\delta))$ the relative norm function from $Q(w)$ to $Q(\delta)$ is

$$N_\delta(\alpha+\beta w) = \alpha^2 - \beta^2 w^2 = \alpha^2 - \beta^2(1+\delta)$$

and we therefore have the obvious definition $N_\delta(A) = N_\delta(\alpha+\beta w)$ for $A = \alpha(1,1) + \beta W \in M(w)$. Note that $AA^* = N_\delta(A)(1,1) = N_\delta(A^*)(1,1)$.

TABLE 2.2 Norms of relative minima of $M(\sqrt{1+\sqrt{-10}})$

k	0	1	2	3	4	5	6	7
$N_\delta(A_k)$	1	$-3-\delta$	5	-2	5	$-3-\delta$	1	$-3-\delta$

Thus all the relative minima listed in table 2.1 have small norms and in particular A_6 corresponds to a unit of Q(w). Furthermore it is easily checked that $A_7 = A_6A_1$ and in fact since M(w) is closed under multiplication we have the more general result

A_k a relative minimum of M(w)

implies (2)

$A_kA_6^j$, $j \in Z$ is a relative minimum of M(w)

To see this note that if $A_kA_6^j$ is not a relative minimum then $\exists\ A \in M(w)\backslash\{(0,0)\}$ such that

$$\text{either} \quad |A| < |A_kA_6^j| \text{ and } |A^*| \leq |(A_kA_6^j)^*|$$

$$\text{or} \quad |A| = |A_kA_6^j| \text{ and } |A^*| < |(A_kA_6^j)^*|$$

However we would therefore have $AA_6^{-j} \in M(w)$ (A_6 is a unit of M(w), $A_6^{-1} = A_6^*$) and the existence of such an element in M(w) is fairly easily seen to contradict the fact that A_k is a relative minimum.

A simple consequence of (2) is that a complete chain of relative minima of M(w) is given by

$$A_{k+6j} = A_kA_6^j, \quad k = 0,1,\ldots,5 \text{ and } j \in Z$$

Since $N_\delta(A_kA_6^j) = N_\delta(A_k)$ it follows that all relative minima of M(w) have small norms.

We can also use (2) to partially explain the reflective nature of the $N_\delta(A_k)$, $k = 0,1,\ldots,6$. Recall that $|A_{k+1}^*| < |A_k^*|$ and so

$$|A_0| = 1 = |A_6^*A_6| < |A_5^*A_6| < \ldots < |A_1^*A_6| < |A_0^*A_6| = |A_6|$$

Since each $A_k^* A_6$ is a relative minimum we must have

$$A_k^* A_6 \sim A_{6-k}, \qquad k = 0,1,\ldots,6$$

We therefore have

$$|N_\delta(A_k)| = |N_\delta(A_{6-k})|, \qquad k = 0,1,\ldots,6$$

since $N_\delta(A_k) = N_\delta(A_k^*)$, $|N_\delta(A)| = |N_\delta(B)|$ when $A \sim B$, and $N_\delta(A_6) = 1$.

These results are typical of the case $w = \sqrt{\alpha}$, $\alpha \in Q(\delta)$ which will be considered in greater detail in section four of this chapter. //

We finish this introductory section with two further theorems. The first gives the obvious generalization of (5) in chapter one section three.

THEOREM 2.9

(a) M(w) has a finite complete chain
if and only if

$$w \in Q(\delta)\backslash\{0\}$$

(b) If $w \in Q(\delta)\backslash\{0\}$ and $A_0, A_1, \ldots, A_n$ is a half chain then $|A_n^*| = 0$ and $w = \alpha_n/\beta_n$.

(c) If $w \notin Q(\delta)$ then as $k \to \infty$ we have $|A_k| \to \infty$, $|A_k^*| \to 0$, and $\alpha_k/\beta_k \to w$.

PROOF

(a) Suppose $w \in Q(\delta)\backslash\{0\}$. Then $w = \alpha/\beta$ for some $\alpha,\beta \in Z(\delta)$. Let $A = \alpha(1,1) + \beta W$. Since $|A^*| = 0$ we see that any relative minimum of M(w) must satisfy $|A_k|, |A_k^*| \leq |A|$. From theorem 2.1b it is now clear that M(w) can only have finitely many relative minima.

Now suppose $w \notin Q(\delta)$. Note that in this case $|AA^*| = 0$ if and only if $A = (0,0)$. Consequently if A is a relative minimum then $|A|, |A^*| \neq 0$. To prove that M(w) has an infinite chain of relative minima it suffices to prove that $\forall\ r > 0$, $\exists\ A \in M(w)$ satisfying $|A^*| < r$ and A is a relative minimum.

Let

$$B_\beta = \{\beta w\}(1,1) + \beta W, \quad \beta \in Z(\delta)$$

where $\{\beta w\}$ denotes the nearest $Z(\delta)$ integer to βw. Clearly $B_\beta \in M(w)$ and $|B_\beta^*| \leq |1+\delta|/2$. Since $Z(\delta)$ is an infinite set it now follows by a standard argument that we can find $B = B_\beta$ for some $\beta \in Z(\delta)$ satisfying $B \neq (0,0)$, $|B^*| < r$. Now choose $A \in M(w)\backslash\{(0,0)\}$ satisfying $|A^*| \leq |B^*|, |A| \leq |B|, |AA^*|$ minimal. It is easily seen that A is a relative minimum. Since $|A^*| < r$ the proof is complete.

(b) and (c) now follow easily. //

The final result of this section shows that if $w \in \mathbb{R}\backslash\{0\}$ then the relative minima of $M(w) = Z(\delta)[(1,1),W]$ and the relative minima of $M(w) = Z[(1,1),W]$ (see chapter one section four) are essentially the same.

THEOREM 2.10

Let $w \in \mathbb{R}\backslash\{0\}$ and let A_k be any relative minimum of M(w). Then up to a $Z(\delta)$ root of unity factor we have

$$A_k = m(1,1) + nW, \quad m,n \in Z$$

Consequently a half chain of relative minima of M(w) is given by A_0 plus the positive relative minima described in theorem 1.12.

PROOF

The final statement follows easily from chapter one section four

once we have shown A_k is of the stated form. Note that without loss of generality we can assume that for $d = -3$ we have

$$\arg(\alpha_k+\beta_k w) \in (-\pi/3,\pi/3],\ \arg(\alpha_k-\beta_k w) \in (-\pi/3,\pi/3] \cup (2\pi/3,4\pi/3]$$

This is because $Z(\sqrt{-3})$ contains the 6 sixth roots of unity $((1+\sqrt{-3})/2)^j$, $j = 0,1,\ldots,5$ and $A_k \in E_k$ implies $\bar{A}_k \in E_k$. Let

$$\alpha_k = (a+b\delta)/c,\quad \beta_k = (e+f\delta)/c$$

Of course if $b = f = 0$ then A_k is clearly of the required form. It therefore remains to consider the case where b,f are not both zero.

In such a case we have

$$B = b(1,1) + fW \in M(w)\setminus\{(0,0)\}$$

Note that

$$|A_k| = |\alpha_k+\beta_k w| = |(a+ew)/c + (b+fw)\delta/c|$$

$$|A_k^*| = |\alpha_k-\beta_k w| = |(a-ew)/c + (b-fw)\delta/c|$$

We now consider two cases.

<u>$d \neq -3$</u> We have $|\delta|/c \geq 1$, $\delta = |\delta|i$ and so

$$|B| = |b+fw| \leq |(b+fw)\delta/c| \leq |A_k|$$

$$|B^*| = |b-fw| \leq |(b-fw)\delta/c| \leq |A_k^*|$$

Since A_k is a relative minimum we have must $|B| = |A_k|$, $|B^*| = |A_k^*|$ which can only occur when $|\delta| = 1$, $a \pm ew = 0$, that is $\delta = \sqrt{-1}$, $a = e = 0$. Thus if b,f are not both zero we conclude that

$$A_k = \delta(b(1,1) + fw),\quad \delta = \sqrt{-1}$$

Since $\sqrt{-1}$ is a root of unity A_k is of the required form.

d = -3 We have c = 2. The conditions satisfied by $\arg(\alpha_k \pm \beta_k w)$ imply

$$|(b+fw)\delta/2| \leq \sqrt{3}|A_k|/2, \quad \text{that is } |B| \leq |A_k|$$

$$|(b-fw)\delta/2| \leq \sqrt{3}|A_k^*|/2, \quad \text{that is } |B^*| \leq |A_k^*|$$

Since A_k is a relative minimum we must have $|B| = |A_k|$, $|B^*| = |A_k^*|$ which can only occur if $\arg(\alpha_k+\beta_k w) = \pi/3$, $\arg(\alpha_k-\beta_k w) \in \{\pi/3, 4\pi/3\}$. It then follows easily that $\arg(\alpha_k)$, $\arg(\beta_k) \in \{\pi/3, 4\pi/3\}$ and therefore $\alpha_k = m(1+\sqrt{-3})/2$, $\beta_k = n(1+\sqrt{-3})/2$ with $m,n \in Z$. Thus

$$A_k = ((1+\sqrt{-3})/2)(m(1,1) + nW)$$

is of the required form. //

SECTION TWO

ORDER OF APPROXIMATION OF A RELATIVE MINIMUM

One of the fundamental results from simple continued fraction theory is

$$|p_k - q_k x| \leq 1/q_{k+1}$$

(See (7) in chapter one section three). In this section we partially develop a similar result for relative minima of M(w). Fjellstedt [1953] has shown that for $w \notin Q(\delta)$ the inequality

$$|\alpha - \beta w| < (1+|d|)/|\beta| \qquad \alpha, \beta \in Z(\delta)$$

has infinitely many solutions. Consequently we might expect to be able to develop a result of the form

$$|A_k^*| = |\alpha_k - \beta_k w| < r|d|/|\beta_k|, \quad k > 0$$

for some $r \in \mathbb{R}^+$. In fact we shall prove

$$|A_k^*| < 2\sqrt{2}|\delta|/c\sqrt[4]{3}\pi^{\frac{1}{2}}|\beta| + (9|\delta|)^2/|\beta|^2, \quad k > 0 \qquad (3)$$

where $|\beta| = \max\{|\beta_k|, |\beta_{k+1}|\}$. (Usually $|\beta_{k+1}| \geq |\beta_k|$. However example 2.4 will show that we can have $|\beta_{k+1}| < |\beta_k|$. Note that if β_{k+1} does not exist then we take $|\beta| = |\beta_k|$). In section four of this chapter we use this partial result to derive the bound

$$|A_k^*| < 2\sqrt{2}|\delta|/c\sqrt[4]{3}\pi^{\frac{1}{2}}(|\beta_k| - 1/2), \quad k > 0 \tag{4}$$

The large size of the final term of (3) is of little consequence since we shall only require an asymptotic result when developing bound (4). Consequently the results in this section are by no means best possible. (Sharper results are obtainable by refinement of the proofs. However the extra effort is not warranted since these sharper results do nothing to simplify the development of (4)).

The result in (3) is based on the next two theorems.

THEOREM 2.11

For $r \in \mathbb{R}^+$ let $n(r)$ be the number of $\beta \in Z(\delta)$ satisfying $|\beta| \leq r$. Then

$$n(r) > c\pi r^2/|\delta| - c2\sqrt{2}\pi r$$

PROOF

We begin with the case $d \equiv 2,3 \pmod 4$. Thus $c = 1$. Let n_1 denote the number of $\beta \in Z(\delta) \cap S$ where

$$S = \{x+yi : x,y > 0, \sqrt{x^2+y^2} \leq r\}$$

We have $n(r) > 4n_1$. (Note that S does not contain any real or pure imaginary numbers). For $\beta = a+b\delta \in Z(\delta)$ define

$$S_\beta = \{x+yi : a-1 < x \leq a, (b-1)|\delta| < y \leq b|\delta|\}$$

Note that the S_β form a disjoint covering of the complex plane. Thus the region $R_1 = \bigcup_{\beta \in S} S_\beta$ has area $n_1|\delta|$. Now consider the region

$$R_2 = \{x+yi : \sqrt{x^2+y^2} \leq r - |1+\delta|,\ x,y > 0\}$$

Suppose $x+yi \in R_2$. Then $x+yi$ also lies in $S_{a+b\delta}$ where

$$a-1 < x \leq a,\ b-1 < y/|\delta| \leq b,\ (a,b \in Z^+)$$

We have $|a+b\delta| < |x+(y/|\delta|)\delta| + |1+\delta| \leq r$ and so $a+b\delta \in S$. Consequently $R_1 \supseteq R_2$ and the corresponding area relationship

$$n_1|\delta| \geq \pi(r-|1+\delta|)^2/4 > \pi r^2/4 - 2\pi r|1+\delta|/4$$

gives us the bound for $n(r)$ since $|1+\delta|/|\delta| \leq \sqrt{2}$.

For $d \equiv 1 \pmod 4$ we have $c = 2$. Note that $a+b\delta \in Z(\delta) \cap S$ $(a,b \in Z)$ implies $((2a-1)+(2b-1)\delta)/2 \in Z(\delta) \cap S$. Consequently $n(r) > 8n_1$ where n_1 is the number of $\beta = a+b\delta \in S$ $(a,b \in Z)$. The result now follows by using the lower bound for n_1 derived in the previous paragraph. //

THEOREM 2.12

Let

$$FR(\delta) = \{x+yi : 0 \leq x \leq 1,\ 0 \leq y \leq |\delta|\}$$

(FR standing for fundamental rectangle). Suppose that we have n complex numbers z_j, $j = 1,2,\ldots,n$ contained in $FR(\delta)$. Then for $n \geq 2$ we can find z_j, z_k with $j \neq k$ satisfying

$$|z_j - z_k| < \sqrt{2|\delta|/(\sqrt{3}n-2)} + 2(1+|\delta|)/(\sqrt{3}n-2)$$

PROOF

The result is trivially true if two of the n complex numbers coincide. Therefore assume $z_j \neq z_k$ when $j \neq k$. Centre a circular region C_j of radius s on each z_j with s chosen so that at least two C_j touch but no two C_j overlap. (C_j, C_k overlap if $C_j \cap C_k$ has non-zero area). We can therefore clearly find z_j, z_k satisfying $|z_j - z_k| = 2s$. Let

$$FR(\delta)_s = \{x+yi : -s \le x \le 1+s,\ -s \le y \le |\delta| + s\}$$

Now the n circular regions C_j are completely contained within $FR(\delta)_s$ but do not occupy the entire area. Let

$$t = (\text{area } \bigcup_{j=1}^{n} C_j) / (\text{area } FR(\delta)_s) = n\pi s^2/(1+2s)(|\delta| + 2s)$$

Clearly $t < 1$ but we can obtain a better bound as follows. The rectangle $FR(\delta)_s$ and the circular regions C_j form a pattern P which (when we consider all possible translations of P of the form

$$P + \ell(1+2s) + im(|\delta|+2s), \quad \ell, m \in Z)$$

defines a packing of the plane by equal radius non-overlapping circles with a corresponding packing constant of t. (See chapter one section one). Consequently $t \le \pi/2\sqrt{3}$ and so we have

$$n\pi s^2/(1+2s)(|\delta|+2s) \le \pi/2\sqrt{3}$$

Rearranging this inequality gives

$$s^2(2\sqrt{3}n-4) - 2s(1+|\delta|) - |\delta| \le 0$$

which implies

$$s \le \frac{2(1+|\delta|) + \sqrt{4(1+|\delta|)^2 + 4|\delta|(2\sqrt{3}n-4)}}{2(2\sqrt{3}n-4)}$$

$$< \frac{(1+|\delta|) + (1+|\delta|) + \sqrt{|\delta|(2\sqrt{3}n-4)}}{(2\sqrt{3}n-4)}$$

$$= \sqrt{|\delta|/(2\sqrt{3}n-4)} + (1+|\delta|)/(\sqrt{3}n-2) \qquad //$$

We now use theorems 2.11, 2.12 to prove the main result of this section.

THEOREM 2.13

Let A_k, $k > 0$ be a relative minimum of M(w) and let $|\beta| = \max\{|\beta_k|, |\beta_{k+1}|\}$. (See the first paragraph of this section).

Then

$$|A_k^*| < 2\sqrt{2}|\delta|/c\sqrt[4]{3}\,\pi^{\frac{1}{2}}|\beta| + (9|\delta|)^2/|\beta|^2$$

PROOF

We begin with the $d \equiv 2,3 \pmod 4$. Thus $c = 1$.

If A_{k+1} does not exist then $|A_k^*| = 0$ and the result is trivial. Therefore assume A_{k+1} exists. Since $|A_k^*| < 1$ for $k > 0$ the theorem is trivially true if $|\beta| \le 9|\delta|$. Therefore assume $|\beta| > 9|\delta|$. Let $r = (|\beta|-1)/2$. For $\gamma \in Z(\delta)$ satisfying $|\gamma| \le r$ let $x_\gamma = (\eta_\gamma - \gamma w)$, with $\eta_\gamma \in Z(\delta)$ chosen so that $x_\gamma \in FR(\delta)$. This gives $n(r)$ complex numbers in $FR(\delta)$ and by theorem 2.12 we can find two numbers $x_{\gamma_1}, x_{\gamma_2}$ satisfying

$$|x_{\gamma_1} - x_{\gamma_2}| < \sqrt{2|\delta|/(\sqrt{3}\,n(r)-2)} + 2(1+|\delta|)/(\sqrt{3}\,n(r)-2)$$

Let $\eta = \eta_{\gamma_1} - \eta_{\gamma_2}$, $\gamma = \gamma_1 - \gamma_2$, $A = \eta(1,1) + \gamma W$. Note that $A \in M(w)\backslash\{(0,0)\}$ since $\gamma_1 \neq \gamma_2$. Furthermore

$$|A^*| = |\eta - \gamma w| = |x_{\gamma_1} - x_{\gamma_2}|$$

We now show that $|A_k^*| \le |A^*|$. If $|A^*| \ge |A_0|$ then this result is trivial since $|A_k^*| < |A_0|$. Therefore suppose $|A^*| < |A_0|$. Note that $|\gamma| \le |\gamma_1| + |\gamma_2| \le |\beta| - 1$ and so

$$\begin{aligned} |A| &\le 2|\gamma w| + |A^*| && \text{by theorem 2.1a} \\ &\le 2|\beta w| - 2|w| + |A^*| \\ &< 2|\beta w| - |A_0| && \text{since } |A^*| < |A_0| \le |w| \end{aligned}$$

Now if $|\beta| = |\beta_k|$ then

$$\begin{aligned} |A| &< 2|\beta_k w| - |A_k^*| && \text{since } |A_k^*| < |A_0| \\ &\le |A_k| && \text{by theorem 2.1a} \\ &< |A_{k+1}| \end{aligned}$$

If however $|\beta| = |\beta_{k+1}|$ then

$$|A| < 2|\beta_{k+1}w| - |A^*_{k+1}| \leq |A_{k+1}|$$

In either case $|A| < |A_{k+1}|$. Now if $|A^*| < |A^*_k|$ then either A or some $B \in M(w)$ satisfying $|A_k| < |B| \leq |A|$, $|B^*| < |A^*_k|$ will be a relative minimum of M(w). However this contradicts the fact that A_{k+1} is a relative minimum of minimal magnitude greater than $|A_k|$. Therefore we must have $|A^*_k| \leq |A^*|$. Thus to this point we have the bound

$$|A^*_k| < \sqrt{2|\delta|/(\sqrt{3}n(r)-2)} + 2(1+|\delta|)/(\sqrt{3}n(r)-2) \qquad (5)$$

We now use theorem 2.11 to obtain the bound stated in the theorem. From theorem 2.11 we have

$$\sqrt{3}n(r)-2 > \sqrt{3}\pi r^2/|\delta| - 2\sqrt{6}\pi r - 2$$

Substituting $r = (|\beta|-1)/2$ gives

$$\sqrt{3}n(r)-2 > \frac{\sqrt{3}\pi}{4|\delta|}(|\beta|^2-2|\beta|+1) - \sqrt{6}\pi(|\beta|-1) - 2$$

$$= \frac{\sqrt{3}\pi|\beta|^2}{4|\delta|} - \sqrt{3}\pi|\beta|\left(\frac{1}{2|\delta|}+\sqrt{2}\right) + \frac{\sqrt{3}\pi}{4|\delta|} + \sqrt{6}\pi - 2$$

$$> \frac{\sqrt{3}\pi|\beta|^2}{4|\delta|} - \sqrt{3}\pi|\beta|\left(\frac{1+2\sqrt{2}}{2}\right)$$

$$= \frac{\sqrt{3}\pi|\beta|^2}{4|\delta|}\left(1 - \frac{2(1+2\sqrt{2})|\delta|}{|\beta|}\right)$$

$$> \frac{\sqrt{3}\pi|\beta|^2}{4|\delta|}\left(1 - \frac{8|\delta|}{|\beta|}\right)$$

Note that the final expression is positive for $|\beta| > 9|\delta|$. Therefore

$$\frac{1}{\sqrt{3}n(r)-2} < \frac{4|\delta|}{\sqrt{3}\pi|\beta|^2}\left(1 - \frac{8|\delta|}{|\beta|}\right)^{-1}$$

$$= \frac{4|\delta|}{\sqrt{3}\pi|\beta|^2}\left(1 + \frac{8|\delta|}{|\beta|}\left(1 - \frac{8|\delta|}{|\beta|}\right)^{-1}\right)$$

$$< \frac{4|\delta|}{\sqrt{3}\pi|\beta|^2}\left(1+\frac{72|\delta|}{|\beta|}\right) \qquad \text{since } |\beta| > 9|\delta|$$

Substituting this result in (5) gives

$$|A_k^*| < \sqrt{\frac{8|\delta|^2}{\sqrt{3}\pi|\beta|^2}\left(1+\frac{72|\delta|}{|\beta|}\right)} + \frac{2(1+|\delta|)4|\delta|}{\sqrt{3}\pi|\beta|^2}\left(1+\frac{72|\delta|}{|\beta|}\right)$$

Now $\sqrt{(1+72|\delta|/|\beta|)} < 1+36|\delta|/|\beta|$, $1+|\delta| \le 2|\delta|$, and $(1+72|\delta|/|\beta|) < 9$ for $|\beta| > 9|\delta|$. Therefore

$$|A_k^*| < \frac{2\sqrt{2}|\delta|}{\sqrt[4]{3}\pi^{\frac{1}{2}}|\beta|}\left(1+\frac{36|\delta|}{|\beta|}\right)+\frac{144|\delta|^2}{\sqrt{3}\pi|\beta|^2}$$

$$= \frac{2\sqrt{2}|\delta|}{\sqrt[4]{3}\pi^{\frac{1}{2}}|\beta|}+\left(\frac{72\sqrt{2}}{\sqrt[4]{3}\pi^{\frac{1}{2}}}+\frac{144}{\sqrt{3}\pi}\right)\frac{|\delta|^2}{|\beta|^2}$$

Since $72\sqrt{2}/\sqrt[4]{3}\pi^{\frac{1}{2}} + 144/\sqrt{3}\pi < 9^2$ the result is now clear.

The proof of the case $d \equiv 1 \pmod 4$ is similar to the above case. However one important difference must be noted. For $d \equiv 1 \pmod 4$ there are at least two choices for η_γ. One choice is of the form $a + b\delta$, $a,b \in Z$ and another choice is of the form $(a+b\delta)/2$, $a,b \in Z$. This gives $cn(r) = 2n(r)$ points in $FR(\delta)$ as opposed to only $n(r)$ points in the case $d \equiv 2,3 \pmod 4$. This difference plus the factor $c = 2$ in theorem 2.11 account for the factor $c = 2$ in the final result. //

SECTION THREE

AN ALGORITHM FOR THE CALCULATION OF RELATIVE MINIMA

In this section we develop an algorithm for calculating a chain of relative minima $A_0, A_1, A_2, \ldots$ in $M(w)$ which is based on the idea of algorithm 1.4. The algorithm we develop suffers the same drawback as the general simple continued fraction algorithm (algorithm 1.1) in that its use in practice is limited by the precision to which calculations can be performed. Of course in practice we would only use the algorithm

developed in this section for $w \notin \mathbb{R}$. For $w \in \mathbb{R}$ the simple continued fraction algorithm (see theorem 2.10) is more appropriate.

We begin with the definition of several types of objects which are at the heart of the algorithm to be developed.

DEFINITION 2.8

Let A_k be a relative minimum of $M(w)$ satisfying $|A_k A_k^*| \neq 0$. We define

$$M_k(w) = M(w)/A_k$$

and for all $j \in Z$

$$E_{k,j} = \{A_j/A_k : A_j \in E_j\}$$

If $A_k^{(n)} \sim A_k$ then $M_k^{(n)}(w)$ will have the obvious meaning. We shall use $R_{k,j}$ to denote an arbitrary element of $E_{k,j}$ and where we wish to denote more than one such element we shall use the notation $R_{k,j}, R_{k,j}^{(1)}, \ldots$ Finally if $j = k+1$ then we shall use the briefer form R_k for $R_{k,k+1}$. //

Given $A_k, A_k^{(1)}$ it can happen that $M_k(w)$ and $M_k^{(1)}(w)$ are not the same sets. (See example 2.5b where $M_3(w) \neq M_3^{(1)}(w)$). If $A_k = UA_k^{(1)}$, $U = (\xi,\xi)$ with ξ a $Z(\delta)$ root of unity, then we will indeed have $M_k(w) = M_k^{(1)}(w)$. Note however that $M_k(w)$ and $M_k^{(1)}(w)$ are always isomorphic ($A/A_k \in M_k(w)$ corresponds to $A/A_k^{(1)} \in M_k^{(1)}(w)$) and that $|A/A_k| = |A/A_k^{(1)}|$, $|(A/A_k)^*| = |(A/A_k^{(1)})^*|$. Therefore the relative minima structures of $M_k(w)$ and $M_k^{(1)}(w)$ are identical.

The following result is obvious yet important. It is obvious in that it is really just a restatement of special cases of theorems 2.2, 2.8 using the notation of definitions 2.7, 2.8.

THEOREM 2.14

A complete chain of relative minima of $M_k(w)$ is given by

$$\ldots, R_{k,-1}, R_{k,0}, R_{k,1}, \ldots$$

In particular $(1,1) = A_k/A_k \sim R_{k,k}$, and $R_k = A_{k+1}/A_k$ satisfies

(a) $|R_k^*| < 1$

(b) $\forall\, S \in M_k(w)\backslash\{(0,0)\}$ satisfying $|S^*| < 1$

we have either $|S| > |R_k|$

or $|S| = |R_k|$ and $|S^*| \geq |R_k^*|$

Finally, given the chain $A_0, A_1, \ldots, A_k$ plus R_k we can extend the chain by taking $A_{k+1} = A_k R_k$. //

The final statement of theorem 2.14 reflects the basic idea for the algorithm we will develop. That is (as in algorithm 1.4) we will calculate a chain $A_0, A_1, A_2, \ldots$ in $M(w)$ indirectly by calculating $R_k \in M_k(w)$, $k = 0,1,2,\ldots$ Now this basic idea is of little use by itself since it suggests nothing more than a scaling of $M(w)$ and a change of notation. Of course the crucial fact is that we will be able to choose a representation of $M_k(w)$ (not $M_k(w) = Z(\delta)[(1,1)/A_k, W/A_k]$) which results in the amount of work required to locate R_k being $O(|\delta|)$. (In example 2.1 the amount of work required to find A_k is $O(|\beta_k|^2)$). Therefore the next step in the development of our algorithm is to develop the appropriate representation of $M_k(w)$.

Ideally we would like to be able to represent $M_k(w)$ in the form

$$M_k(w) = Z(\delta)[(1,1), W_k]$$

for some $W_k \in M_k(w)$. (The reason why such a representation is desirable

will become clearer when we look at the actual calculation of R_k). Now (1,1) belongs to a basis of $M_k(w)$ if and only if A_k belongs to a basis of M(w) which occurs if and only if $\alpha_k\lambda - \beta_k\kappa = 1$ for some $\kappa,\lambda \in Z(\delta)$. (See comments following definitions 2.2, 2.3. If κ,λ exist then for $B = \kappa(1,1) + \lambda W$ we have $M(w) = Z(\delta)[A_k,B]$ and so $M_k(w) = Z(\delta)[(1,1),B/A_k]$, that is we can take $W_k = B/A_k$). However the following example shows that a relative minimum need not belong to a basis of M(w). (Note that if $Z(\delta)$ is a unique factorization domain then it is easily seen that a relative minimum must belong to a basis of M(w)).

EXAMPLE 2.2

In example 2.1 we saw that $A_3 = (2+5\delta)(1,1) + (6+2\delta)W$ is a relative minimum of M(w), $w = \sqrt{1+\delta}$, $\delta = \sqrt{-10}$. Now if A_3 belongs to a basis of M(w) then we will be able to find $\kappa,\lambda \in Z(\delta)$ for which

$$(2+5\delta)\lambda - (6+2\delta)\kappa = 1 \qquad (6)$$

Let $\kappa = a + b\delta$, $\lambda = e + f\delta$ with $a,b,e,f \in Z$. Substituting for κ,λ in (6) and then separating out the real part gives the equation

$$2e - 50f - 6a + 20b = 1$$

which clearly has no solution. Consequently (6) has no solution and we can conclude that A_3 does not belong to a basis of M(w). //

Clearly a more general form for representing $M_k(w)$ is required. To develop such a representation we shall need the following theorem and definition.

THEOREM 2.15

Let α_k,β_k be the coefficients of a relative minimum. Then α_k,β_k have no common rational integer factor and so

$$<\alpha_k,\beta_k> = Z[g,\sigma]$$

where g is the minimal positive rational integer in $\langle\alpha_k,\beta_k\rangle$ and $\sigma = (a+\delta)/c \in Z(\delta)$ with $0 \leq a < cg$. Furthermore

$$g = (N(\alpha_k), N(\beta_k), N(\alpha_k+\beta_k))$$

PROOF

Since α_k,β_k are the coefficients of a relative minimum they clearly have no common $Z(\delta)$ factor and so the form of $\langle\alpha_k,\beta_k\rangle$ follows from theorems 1.4, 1.5. Therefore the only point which really requires explanation is the final one. Let $f = (N(\alpha_k), N(\beta_k), N(\alpha_k+\beta_k))$. Now as $g \in \langle\alpha_k,\beta_k\rangle$ we have $g = \alpha_k\alpha + \beta_k\beta$ where $\alpha,\beta \in Z(\delta)$. Thus

$$g\alpha_k' = N(\alpha_k)\alpha + \alpha_k'\beta_k\beta, \quad g\beta_k' = \alpha_k\beta_k'\alpha + N(\beta_k)\beta$$

By theorem 1.2 we have $f|\alpha_k\beta_k', \alpha_k'\beta_k$ and so $f|g\alpha_k', g\beta_k'$. Since α_k', β_k' have no common rational integer factor we conclude that $f|g$. However since we can find $\ell,m,n \in Z$ such that

$$\begin{aligned} f &= \ell N(\alpha_k) + mN(\beta_k) + nN(\alpha_k+\beta_k) \\ &= \alpha_k(\ell\alpha_k' + n(\alpha_k'+\beta_k')) + \beta_k(m\beta_k' + n(\alpha_k'+\beta_k')) \end{aligned}$$

we have $f \in \langle\alpha_k,\beta_k\rangle$ and so $g|f$. Thus $g = f$. //

DEFINITION 2.9

A relative minimum $A_k \in M(w)$ will be called *basic (non-basic)* if A_k belongs (does not belong) to a basis of $M(w)$.

For A_k a relative minimum of $M(w)$ we define

$$I_k = \langle\alpha_k,\beta_k\rangle$$

The standard representation of I_k (see theorems 2.15, 1.4) will be

$$I_k = Z[g_k,\sigma_k], \quad \sigma_k = (a_k+\delta)/c$$

We shall use κ_k, λ_k to denote $Z(\delta)$ integers satisfying

$$\alpha_k\lambda_k - \beta_k\kappa_k = g_k$$

and finally we define

$$B_k = \kappa_k(1,1) + \lambda_k W, \quad W_k = B_k/A_k \qquad //$$

Note that A_k is basic if and only if $g_k = 1$. If A_k is basic then $M(w) = Z(\delta)[A_k, B_k]$ and $M_k(w) = Z(\delta)[(1,1), W_k]$. Of course κ_k, λ_k (and hence B_k, W_k) are not unique since

$$\alpha_k(\lambda_k+\theta\beta_k) - \beta_k(\kappa_k+\theta\alpha_k) = g_k, \quad \forall\ \theta \in Z(\delta)$$

(In general there are other solutions as well as those given).

We now give a standard representation of $M_k(w)$ which will be used in the algorithm. (Other non-standard representations are possible if we allow $\alpha_k\lambda_k - \beta_k\kappa_k = \xi g_k$ where $\xi \neq 1$ is a $Z(\delta)$ root of unity. We shall assume that $M_k(w)$ representations are of the following standard form unless otherwise stated).

THEOREM 2.16

Suppose A_k is a relative minimum of $M(w)$ and $I_k, \kappa_k, \lambda_k, W_k$ are as defined in definition 2.9. A standard representation of $M_k(w)$ is given by

$$M_k(w) = \{(\alpha(1,1) + \beta W_k)/g_k : (\alpha,\beta)\ W_k \text{ allowable}\}$$

By (α,β) W_k allowable we mean

$$\alpha = \kappa\lambda_k - \lambda\kappa_k, \quad \beta = -\kappa\beta_k + \lambda\alpha_k, \quad \kappa,\lambda \in Z(\delta) \qquad (7)$$

Furthermore we can find a unique $\psi_k \in Z$ (for the given W_k) satisfying $0 \leq \psi_k < g_k$ such that

(α,β) W_k allowable

if and only if

$$(\alpha,\beta) = m_1(g_k,0) + m_2(\sigma_k',0) + m_3(0,g_k) + m_4(\psi_k,\sigma_k),\ m_j \in Z$$

(The terminology (α,β) W_k allowable reflects the fact that α,β are allowable as coefficients of an element of $M_k(w)$. (For the given W_k). Note that if A_k is basic then $g_k = 1$, $\psi_k = 0$, $\sigma_k = \omega$ and $M_k(w) = Z(\delta)[(1,1),W_k]$. However if A_k is non-basic then $(1,1)$ does not belong to a basis of $M_k(w)$ and so the representation of $M_k(w)$ is not in terms of a module basis. Finally we note that it is easily shown that $\psi_k = 0$ if and only if $\kappa_k, \lambda_k \in I_k'$. Such κ_k, λ_k can always be found (see the proof of theorem 2.15) and it is therefore possible to have a simpler representation of $M_k(w)$. However this simpler representation is usually more difficult to obtain especially if we encounter a chain of two or more non-basic relative minima and so we have not pursued this matter).

PROOF

It is easily checked that

$$(1,1) = (\lambda_k A_k - \beta_k B_k)/g_k,\ W = (-\kappa_k A_k + \alpha_k B_k)/g_k$$

Thus

$$M(w) = \{\kappa(1,1) + \lambda W : \kappa, \lambda \in Z(\delta)\}$$

$$= \{((\kappa\lambda_k - \lambda\kappa_k)A_k + (-\kappa\beta_k + \lambda\alpha_k)B_k)/g_k : \kappa,\lambda \in Z(\delta)\}$$

$$= \{(\alpha A_k + \beta B_k)/g_k : (\alpha,\beta)\ W_k \text{ allowable}\}$$

and so

$$M_k(w) = \{(\alpha(1,1) + \beta W_k)/g_k : (\alpha,\beta)\ \ W_k \text{ allowable}\}$$

It therefore remains to show that the W_k allowable (α,β) are precisely those pairs of $Z(\delta)$ integers described in the if and only if condition of the theorem.

Note that $(g_k,0)$, $(0,g_k)$ are W_k allowable. (Take $\kappa=\alpha_k,\lambda=\beta_k$ and $\kappa=\kappa_k,\lambda=\lambda_k$ in (7)). Furthermore with the aid of (7) it is easily checked that if (α,β), (γ,θ) are W_k allowable then so is

$$\kappa(\alpha,\beta)+\lambda(\gamma,\theta) = (\kappa\alpha+\lambda\gamma,\kappa\beta+\lambda\theta) \quad \forall\ \kappa,\lambda \in Z(\delta)$$

The set of W_k allowable β coefficients (that is the set of β for which $\exists\ \alpha \in Z(\delta)$ such that (α,β) is W_k allowable) is clearly $I_k = <\alpha_k,\beta_k> = Z[g_k,\sigma_k]$. Thus any W_k allowable β coefficient is uniquely representable as $m_3g_k+m_4\sigma_k$ where $m_3, m_4 \in Z$. Let $\psi_k = (e+f\delta)/c$ be the $Z(\delta)$ integer satisfying (ψ_k,σ_k) W_k allowable, f non-negative and minimal, and e non-negative and minimal for this value of f. We can therefore write any W_k allowable (α,β) as

$$(\alpha,\beta) = (\alpha-m_4\psi_k,0)+m_3(0,g_k)+m_4(\psi_k,\sigma_k) \tag{8}$$

The set of all α such that $(\alpha,0)$ is W_k allowable is easily seen to be an ideal I. We have

$$\begin{aligned}
\alpha \in I &\iff (\alpha(1,1)+0W_k)/g_k && \in M_k(w)\\
&\iff \alpha A_k/g_k && \in M(w)\\
&\iff (\alpha\alpha_k/g_k)(1,1)+(\alpha\beta_k/g_k)W && \in M(w)\\
&\iff \alpha\alpha_k/g_k,\ \alpha\beta_k/g_k && \in Z(\delta)
\end{aligned}$$

Clearly $g_k \in I$ and in fact g_k must be the minimal positive rational integer in I since α_k, β_k have no common factor. Now $\sigma_k \in <\alpha_k,\beta_k>$ implies $\sigma_k = \kappa\alpha_k+\lambda\beta_k$ for some $\kappa,\lambda \in Z(\delta)$. Theorems 1.2, 2.15 imply $g_k | N(\alpha_k), N(\beta_k), \alpha_k\beta_k', \alpha_k'\beta_k$ and it therefore follows that $\sigma_k'\alpha_k/g_k$, $\sigma_k'\beta_k/g_k \in Z(\delta)$. Thus $\sigma_k' \in I$. It is now not difficult to see that we must have $I = I_k'$.

Returning to (8) we now have that $\alpha - m_4\psi_k = m_1 g_k + m_2\sigma_k'$ for unique $m_1, m_2 \in Z$. We therefore have that any W_k allowable (α,β) is of the form

$$(\alpha,\beta) = m_1(g_k,0) + m_2(\sigma_k',0) + m_3(0,g_k) + m_4(\psi_k,\sigma_k),\ m_j \in Z.$$

Furthermore it is clear that any such (α,β) is W_k allowable since each of $(g_k,0)$, $(\sigma_k',0)$, $(0,g_k)$, (ψ_k,σ_k) is W_k allowable.

To complete the proof we must show that ψ_k satisfies the stated conditions. Since $m_1(g_k,0) + m_2(\sigma_k',0) + (\psi_k,\sigma_k)$ is W_k allowable and $\sigma_k' = (a_k-\delta)/c$ it is clear that $f = 0$ and $0 \le e < cg_k$. That is $\psi_k = e/c \in Z$ and $0 \le \psi_k < g_k$. The uniqueness of ψ_k follows from the fact that g_k is the minimal integer in $I_k' \cap Z^+$.

COROLLARY

$$g_k \le 2|\delta|/c\sqrt{3}$$

PROOF

Since $(\sigma_k'/g_k)(1,1), ((\sigma_k'-g_k)/g_k)(1,1) \in M_k(w)$ we have $R = ((a-\delta)/cg_k)(1,1) \in M_k(w)$ with $|a| \le cg_k/2$. However since $(1,1)$ is a relative minimum of $M_k(w)$ and $|R| = |R^*|$ we must have $|R| \ge 1$. Thus

$$\begin{aligned} |(a-\delta)/cg_k| \ge 1 &\Rightarrow g_k^2c^2 \le a^2 - d \\ &\Rightarrow g_k^2c^2 - g_k^2c^2/4 \le -d \\ &\Rightarrow g_k \le 2|\delta|/c\sqrt{3} \end{aligned}$$ //

Although the representation of $M_k(w)$ in theorem 2.16 may seem a little cumbersome it will prove quite satisfactory when it comes to calculating R_k in $M_k(w)$. We shall not calculate the required representation of $M_k(w)$ from α_k, β_k since these coefficients eventually become unmanageably large. Instead we shall use coefficients which arise in the calculation of $R_{k-1} \in M_{k-1}(w)$. Consequently we delay

discussion of the practical aspects of calculating a representation of $M_k(w)$ until later in this section.

We now consider the problem of calculating $R_k \in M_k(w)$. The following notation will be needed.

DEFINITION 2.10

Suppose $M_k(w)$ is represented as in theorem 2.16. The components of W_k will be denoted by $w_{k,1}, w_{k,2}$ and the coefficients of R_k will be denoted by $\eta_{k,1}, \eta_{k,2}$. That is

$$W_k = (w_{k,1}, w_{k,2}),\ R_k = (\eta_{k,1}(1,1) + \eta_{k,2}W_k)/g_k$$

with $w_{k,1}, w_{k,2} \in \mathbb{C}$ and $(\eta_{k,1}, \eta_{k,2})$ W_k allowable. //

Let us suppose that we have available a standard representation of $M_k(w)$ as described in theorem 2.16 which will be obtained by a process outlined later in this section. Now from theorem 2.14 we see that we can take R_k to be any $R \in M_k(w)\backslash\{(0,0)\}$ which satisfies

(a) $|R^*| < 1$, $|R|$ minimal

(b) $S \in M_k(w)$ and $|S| = |R|$ implies $|S^*| \geq |R^*|$

We shall find the required R_k by using a search of the type used in example 2.1 to find A_1. The next two theorems help to clarify this procedure by indicating the maximum amount of work which may be required by the search.

THEOREM 2.17

Let

$$FR(\delta) = \{x+yi : 0 \leq x \leq 1,\ 0 \leq y \leq |\delta|\}$$

Suppose that we have n complex numbers x_j, $j = 1,2,\ldots,n$ satisfying $x_j \in FR(\delta)$ with

$$n > 2|\delta|/\sqrt{3}c$$

Then we can find x_j, x_m with $1 \leq j,m \leq n$ and $j \neq m$ such that

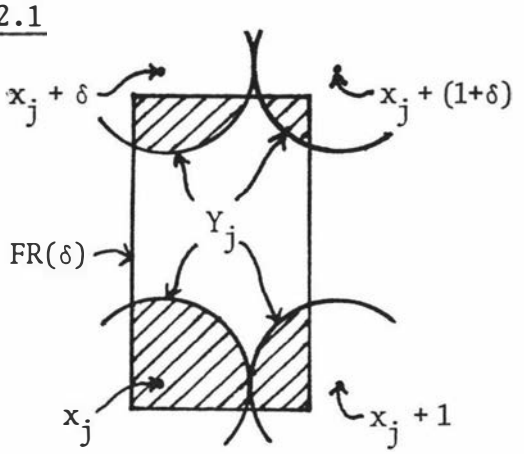
$$|x_j - x_m + \gamma| < 1$$

for some $\gamma \in Z(\delta)$.

PROOF

We begin with the case $d \equiv 2,3 \pmod 4$ and so $c = 1$. Suppose no pair x_j, x_m with $j \neq m$ satisfy the inequality. We show that this leads to a contradiction.

Figure 2.1



For each x_j let

$$X_j = \{x \in \mathbb{C} : |x_j - x| \leq \tfrac{1}{2}\}$$

and let

$$Y_j = \bigcup_{\gamma \in Z(\delta)} ((X_j + \gamma) \cap FR(\delta))$$

where $X_j + \gamma = \{x + \gamma : x \in X_j\}$. (See fig 2.1). Note that $Y_j \subset FR(\delta)$. Furthermore since $|\gamma_1 - \gamma_2| \geq 1$ for $\gamma_1, \gamma_2 \in Z(\delta)$, $\gamma_1 \neq \gamma_2$, the area of $(X_j + \gamma_1) \cap (X_j + \gamma_2)$ must be zero. Consequently it is clear that Y_j and X_j have equal areas of $\pi/4$. Now if any two regions Y_j, Y_m $(j \neq m)$ have a non-zero area of intersection then it is easily seen that $|x_j - x_m + \gamma| < 1$ for some $\gamma \in Z(\delta)$ contrary to our supposition at the beginning of the proof. Therefore the Y_j must have zero area of intersection and so $Y = \bigcup_{j=1}^{n} Y_j$ has area $n\pi/4$. By using the same argument as in the proof of theorem 2.12 we see that Y has area at most $\pi|\delta|/2\sqrt{3}$. Consequently $n\pi/4 \leq \pi|\delta|/2\sqrt{3}$. This implies $n \leq 2|\delta|/\sqrt{3}$ which is a contradiction.

The proof for the case $d \equiv 1 \pmod 4$ is virtually identical to the above proof. The only difference is that Y_j has area $c\pi/4 = \pi/2$

rather than $\pi/4$. The additional area comes from translations of the form $X_j + \gamma$ where $\gamma = (a+b\delta)/2$ with a,b odd.

COROLLARY

We can find $R = (\alpha(1,1) + \beta W_k)/g_k \in M_k(w)\backslash\{(0,0)\}$ satisfying

$$|R^*| < 1, \quad |\beta| \leq 2|\delta|/\sqrt{3}c, \quad \beta \in Z, \quad g_k|\beta.$$

PROOF

Let $\ell = [2|\delta|/\sqrt{3}c\, g_k]$. Note that as $(\theta g_k, 0)$ is W_k allowable $\forall\ \theta \in Z(\delta)$ we have $R \in M_k(w)$ implies $R + \theta(1,1) \in M_k(w), \forall\ \theta \in Z(\delta)$. Therefore for $m = 0,1,\ldots,\ell$ and $n = 0,1,\ldots, g_k - 1$ we can choose $\theta_{m,n} \in Z(\delta)$ such that

$$S_{m,n} = \theta_{m,n}(1,1) + (n\sigma_k'(1,1) + mg_kW_k)/g_k = (S_{m,n,1}, S_{m,n,2})$$

has $S_{m,n,2} \in FR(\delta)$. Now this gives us $(\ell+1)g_k$ numbers in $FR(\delta)$. Since $(\ell+1)g_k > 2|\delta|/\sqrt{3}c$ theorem 2.17 shows that we can find $S_{m,n,2}, S_{p,q,2}$ with $0 \leq m, p \leq \ell$, $0 \leq n,q \leq g_k - 1$ and $m \neq p$ or $n \neq q$ such that

$$|S_{m,n,2} - S_{p,q,2} + \gamma| < 1 \quad \text{for some } \gamma \in Z(\delta)$$

Let

$$R = S_{m,n} - S_{p,q} + \gamma(1,1)$$

Clearly $R \in M_k(w)$, $|R^*| = |S_{m,n,2} - S_{p,q,2} + \gamma| < 1$. It is not difficult to check that $R \neq (0,0)$. Finally note that $R = (\alpha(1,1) + \beta W_k)/g_k$ with

$$|\beta| = |g_k(m-p)| \leq g_k\ell \leq 2|\delta|/\sqrt{3}c \qquad //$$

The R described in the corollary can be found by successively testing the $\beta \in Z^+ \cap I_k$ (that is the multiples of g_k) in order of increasing magnitude for the existence of an $\alpha \in Z(\delta)$ such that (α,β)

is W_k allowable and $|R^*| < 1$. (If more than one α satisfies the condition then we choose an α for which $|R|$ is minimal. However if this minimal value of $|R|$ occurs for more than one α we choose one of these α for which $|R^*|$ is minimal). The corollary guarantees that we will find such an R with $|\beta| \leq 2|\delta|/c\sqrt{3}$. Now this R will not necessarily be R_k (it often is) but it does give a bound for $|\eta_{k,2}|$.

THEOREM 2.18

Suppose $R = (\alpha(1,1) + \beta W_k)/g_k \in M_k(w)\setminus\{(0,0)\}$ satisfies $|R^*| < 1$.

Then

$$\text{(a)} \quad |\eta_{k,2}| < g_k(|R|+1)/|W_k - W_k^*| \tag{9}$$

$$= (|R|+1)|\alpha_k^2 - \beta_k^2 w^2|/2|w|$$

$$\text{(b)} \quad |\eta_{k,2}| < 2|\delta|/c\sqrt{3} + |\alpha_k^2 - \beta_k^2 w^2|/|w|$$

PROOF

We have $R_k = (\eta_{k,1}(1,1) + \eta_{k,2} W_k)/g_k$, $|R_k| \leq |R|$, $|R_k^*| < 1$.

Thus

$$|\eta_{k,2}(W_k - W_k^*)| = g_k|R_k - R_k^*| < g_k(|R|+1)$$

Now $W_k = B_k/A_k$ for some $B_k = \kappa_k(1,1) + \lambda_k W \in M(w)$ with $\alpha_k\lambda_k - \beta_k\kappa_k = g_k$.

Therefore

$$|W_k - W_k^*| = |B_k/A_k - B_k^*/A_k^*|$$

$$= \left|\frac{\kappa_k + \lambda_k w}{\alpha_k + \beta_k w} - \frac{\kappa_k - \lambda_k w}{\alpha_k - \beta_k w}\right|$$

$$= 2\left|\frac{(\alpha_k\lambda_k - \beta_k\kappa_k)w}{\alpha_k^2 - \beta_k^2 w^2}\right|$$

$$= 2g_k|w/(\alpha_k^2 - \beta_k^2 w^2)|$$

Since $w \neq 0$ we have $|W_k - W_k^*| \neq 0$ and so part (a) is clear. (Note that

$M_k(w)$ is only defined for $|A_k A_k^*| = |\alpha_k^2 - \beta_k^2 w^2| \neq 0)$.

To prove part (b) we note that $g_k R = \beta(W_k - W_K^*) + g_k R^*$ and so from part (a) plus the above expression for $|W_k - W_k^*|$ and the fact $|R^*| < 1$ we have

$$|\eta_{k,2}| < |\beta| + g_k(|R^*|+1)/|W_k - W_k^*|$$

$$< |\beta| + |(\alpha_k^2 - \beta_k^2 w^2)/w|$$

By the corollary to theorem 2.17 we see that we can assume that $|\beta| \leq 2|\delta|/\sqrt{3}c$ which gives result (b). //

Thus to find R_k we first find R as indicated following theorem 2.17. This gives a bound (9) for $|\eta_{k,2}|$ and the remaining possibilities (often none) in $Z[g_k,\sigma_k] \cap Z(\delta)^+$ can be checked. (See chapter one section two for $Z(\delta)^+$. Note that if $d = -1$ we can in actual fact restrict our attention to the β satisfying $\arg(\beta) \in [0,\pi/2)$. This is because $Z(\delta)$ contains the roots of unity $\pm i$ as well as ± 1. Similarly when $d = -3$ we can restrict our attention to the β satisfying $\arg(\beta) \in [0,\pi/3)$. To simplify the presentation we shall however ignore this simplification in the following paragraphs and the resulting algorithm. Of course in practice this simplification is easily made and should always be used).

The bound in theorem 2.18b has been given as an indication of the maximum number of tests which may be required to find R_k. In the proof of theorem 2.33 we shall see that

$$|\alpha_k^2 - \beta_k^2 w^2| \leq 4\sqrt{2}\,|\delta||w|/\sqrt[4]{3}\pi^{\frac{1}{2}}c$$

Thus $|\alpha_k^2 - \beta_k^2 w^2|/|w| < 2.4|\delta|/c$ and so we obtain an upper bound of $3.6|\delta|/c$ for $|\eta_{k,2}|$. Of course this bound describes the worst possible case and in practice we usually find $|\eta_{k,2}| \leq |\omega|$.

Recall that following theorem 2.14 we indicated that a representation of $M_k(w)$ involving (1,1) is most desirable. We can now give a very brief indication of why this is so. More general representations of $M_k(w)$ involving linearly independent elements $T, V \in M_k(w)$ are of course possible. We can also develop corresponding generalizations of the last few theorems. The generalization of theorem 2.18b gives a bound depending on $|T|$ and $|T^*|$. Now we would expect the search for R_k to involve the least work when this bound is minimal. This occurs when $|T|$, $|T^*|$ are approximately equal and minimal. However (1,1) is a relative minimum of $M_k(w)$ so we cannot have $|T|, |T^*| < 1$. The choice $T = (1,1)$ is therefore the most obvious and the simplest one. Of course in any given case other choices of T may give a better bound for $|\eta_{k,2}|$. However the effort required to find such a T would normally far outweigh any savings that are gained in the calculation of R_k.

So far we have seen that the location of R_k involves a relatively simple search. We now look more closely at the details of this search. In particular we consider the order in which the β (coefficients of W_k) are searched through, and the method for determining if there exists $\alpha \in Z(\delta)$ such that (α,β) is W_k allowable and

$$|R^*| = |(\alpha(1,1) + \beta W_k^*)/g_k| = |(\alpha + \beta w_{k,2})/g_k| < 1$$

We have already noted that $|\eta_{k,2}| < 3.6|\delta|/c$ and so we can take $b(\beta) = 3.6|\delta|/c$ as the initial upper bound on the magnitude of the $\beta \in Z(\delta)^+ \cap Z[g_k,\sigma_k]$ which need to be considered in the search for R_k. Now as any $\beta \in Z[g_k,\sigma_k]$ is of the form $m_3 g_k + m_4 \sigma_k$, $m_3, m_4 \in Z$ we can search through the appropriate β in the following order. For $m_4 = 0,1,-1,2,-2\ldots$ (while $|m_4||\delta|/c < b(\beta)$) test $\beta = m_3 g_k + m_4 \sigma_k$, $m_3 = \ell, \ell+1, \ldots, n$ where ℓ (resp. n) is the minimal (resp. maximal) rational integer such that the corresponding β satisfies $|\beta| < b(\beta)$,

$\beta \in Z(\delta)^+$. (Note that ℓ,n may not exist in which case there are no values of β corresponding to the current value of m_4 to test). For any β generated in this manner we must test for the existence of an α satisfying the properties listed previously. Now any α for which (α,β) is W_k allowable is of the form

$$\alpha = m_1 g_k + m_2 \sigma_k' + m_4 \psi_k, m_1, m_2 \in Z, m_4 \text{ determined by } \beta$$

We therefore seek $m_1, m_2 \in Z$ such that

$$|(m_1 g_k + m_2 \sigma_k' + m_4 \psi_k + \beta w_{k,2})/g_k| < 1 \tag{10}$$

If such m_1, m_2 exist then we must certainly have

$$|\mathrm{Im}(\alpha+\beta w_{k,2})/g_k| = |(-m_2|\delta|/c + \mathrm{Im}(\beta w_{k,2}))/g_k|$$
$$< 1$$

This implies

$$c(-g_k+\mathrm{Im}(\beta w_{k,2}))/|\delta| < m_2 < c(g_k+\mathrm{Im}(\beta w_{k,2}))/|\delta|$$

and so

$$[c(-g_k+\mathrm{Im}(\beta w_{k,2}))/|\delta|] + 1 \leq m_2 \leq [c(g_k+\mathrm{Im}(\beta w_{k,2}))/|\delta|]$$

Now $2cg_k/|\delta| \leq 4\sqrt{3} \approx 2.3$ and so there are 0,1,2 or 3 possibilities for m_2. For any one of these possibilities we then require $m_1 \in Z$ satisfying (10). Clearly two possibilities for m_1 will need checking. Thus for a given β there are up to 6 pairs m_1, m_2 (corresponding to 6 values of α) which may lead to R satisfying $|R^*| < 1$.

As the search for R_k proceeds we will produce a number of $R \in M_k(w)$ satisfying $|R^*| < 1$. Of course R_k will be amongst these R and it can be sifted out as follows. Since the R indicated in the corollary to theorem 2.17 satisfies

$$|R| = |\beta(W_k - W_k^*)/g_k + R^*| < 2|\delta||w_{k,1} - w_{k,2}|/\sqrt{3}\, c\, g_k + 1$$

we have an initial upper bound for $|R_k|$ of

$$mR = 2|\delta||w_{k,1} - w_{k,2}|/c\sqrt{3}\, g_k + 1$$

Set $mR^* = 1$. Now when the search produces an $R \in M_k(w)$ for which $|R^*| < 1$ we check to see if $|R| < mR$, or $|R| = mR$ and $|R^*| < mR^*$. If either condition is satisfied then we set $T = R$, $mR = |R|$, $mR^* = |R^*|$, and update $b(\beta)$ using (9). We then continue the search. In this way we keep track of the current best possibility for R_k and when the search is completed we will have $R_k = T$.

The discussion in the last few paragraphs tends to give the impression that the calculation of R_k involves far more work than it actually does in practice. Generally speaking we find $g_k = 1$. Consequently for a given β there is usually at most one possibility for m_2 and often there are no m_2 possibilities. On the other hand when $g_k > 1$ (and so more possibilities for m_2 are likely to occur) we are only considering one in every g_k of the $\beta \in Z(\delta)^+$, that is there will be fewer β coefficients to consider in the search. Finally we normally find that the location of the R indicated in the corollary to theorem 2.17 results in the bound $b(\beta)$ being at least halved.

The ideas of the last few paragraphs are now collected together in the form of an algorithm which will eventually become a major part of our algorithm for the calculation of relative minima.

ALGORITHM 2.1A

Assume that we have available a standard representation of $M_k(w)$. (See theorem 2.16). Then R_k can be calculated as follows.

1 Set $b(\beta) = 3.6|\delta|/c$, $mR = 2|\delta||w_{k,1} - w_{k,2}|/\sqrt{3}\, cg_k + 1$, $mR^* = 1$

2 Set $m_3 = 1$, $m_4 = 0$, $\beta = g_k$

3 Calculate $x = \beta w_{k,2} + m_4 \psi_k$

4 Calculate $m_2 = [c(-g_k + Im(x))/|\delta|] + 1$, $m_2^+ = [c(g_k + Im(x))/|\delta|]$

5 If $m_2 > m_2^+$ go to 15

6 Calculate $m_1 = [-(m_2 a_k/c + Re(x))/g_k]$

7 Set $m_1^+ = m_1 + 1$

8 Calculate $r^* = |m_1 + (m_2\sigma_k' + x)/g_k|$

9 If $r^* \geq 1$ go to 13

10 Calculate $r = |m_1 + (m_2\sigma_k' + m_4\psi_k + \beta w_{k,1})/g_k|$

11 If $r > mR$, or $r = mR$ and $r^* \geq mR^*$ go to 13

12 Set $\eta_{k,1} = m_1 g_k + m_2\sigma_k' + m_4\psi_k$, $\eta_{k,2} = \beta$, $mR = r$, $mR^* = r^*$

$b(\beta) = g_k(r+1)/|w_{k,1} - w_{k,2}|$

13 If $m_1 < m_1^+$ then increment m_1 by 1 and go to 8

14 Increment m_2 by 1 and go to 5

15 Increment m_3 by 1

16 Set $\beta = m_3 g_k + m_4\sigma_k$

17 If $|\beta| < b(\beta)$ go to 3

18 If $m_4 > 0$ reset m_4 to $-m_4$, go to 21

19 Reset m_4 to $-m_4 + 1$

20 If $m_4|\delta|/c \geq b(\beta)$ go to 24

21 Set $m_3 = -[m_4 a_k/c\, g_k]$

22 If $m_4 < 0$ and $m_3 g_k + m_4 a_k/c = 0$ then increment m_3 by 1

23 Go to 16

24 $R_k = (\eta_{k,1}(1,1) + \eta_{k,2} W_k)/g_k$

25 Stop //

We now return to the problem of calculating the required representation of $M_k(w)$. The algorithm for calculation of relative minima will obviously start with the module

$$M_0(w) = Z(\delta)[(1,1),W]/A_0 = Z(\delta)[(1,1),W_0]$$

where $W_0 = W$ if $|w| \geq 1$ and $W_0 = -W^{-1}$ if $|w| < 1$. (We must take *negative* W^{-1} in order that the coefficients $\kappa_0 = -1$, $\lambda_0 = 0$ of $B_0 = A_0 W_0$ satisfy $\alpha_0\lambda_0 - \beta_0\kappa_0 = 1$). We also have

$$g_0 = 1,\ \psi_0 = 0,\ \sigma_0 = \omega$$

Thus a standard representation of $M_0(w)$ is fairly trivial to obtain. More generally suppose that we have a standard representation of $M_k(w)$ (theorem 2.16) plus $R_k \in M_k(w)$ and we wish to calculate a standard representation of $M_{k+1}(w)$ where $A_{k+1} = R_k A_k$. We have already indicated that we wish to avoid working with the coefficients α_k, β_k since they grow with k. The following theorem shows how this can be done.

THEOREM 2.19

Suppose that we have a standard representation of $M_k(w)$ and that we have $R_k = (\eta_{k,1}(1,1) + \eta_{k,2}W_k)/g_k \in M_k(w)$ giving $A_{k+1} = R_k A_k$. Then we can find $\eta_{k,3}, \eta_{k,4} \in Z(\delta)$ such that $(\eta_{k,3}, \eta_{k,4})$ is W_k allowable and

$$\eta_{k,1}\eta_{k,4} - \eta_{k,2}\eta_{k,3} = g_k g_{k+1}$$

Let

$$S_k = (\eta_{k,3}(1,1) + \eta_{k,4}W_k)/g_k$$

Then provided $|R_k R_k^*| \neq 0$ (equivalently $|A_{k+1}A_{k+1}^*| \neq 0$) we have a standard representation of $M_{k+1}(w)$ defined by

$$W_{k+1} = S_k/R_k$$

$$I_{k+1} = Z[-\eta_{k,2}, -\sigma_k'\eta_{k,2}/g_k, \eta_{k,1}, (\sigma_k\eta_{k,1} - \psi_k\eta_{k,2})/g_k]$$

$$\psi_{k+1} \equiv (\alpha\eta_{k,4} - \beta\eta_{k,3})/g_k \pmod{I'_{k+1}}$$

where (α,β) is W_k allowable and $(-\alpha\eta_{k,2} + \beta\eta_{k,1})/g_k = \sigma_{k+1}$.

PROOF

Note that

$$A_{k+1} = A_k R_k = (\eta_{k,1} A_k + \eta_{k,2} B_k)/g_k$$

$$= ((\eta_{k,1}\alpha_k + \eta_{k,2}\kappa_k)(1,1) + (\eta_{k,1}\beta_k + \eta_{k,2}\lambda_k)W_k)/g_k$$

Thus $\alpha_{k+1} = (\eta_{k,1}\alpha_k + \eta_{k,2}\kappa_k)/g_k$ and $\beta_{k+1} = (\eta_{k,1}\beta_k + \eta_{k,2}\lambda_k)/g_k$. Let $\kappa_{k+1},\lambda_{k+1} \in Z(\delta)$ satisfy $\alpha_{k+1}\lambda_{k+1} - \beta_{k+1}\kappa_{k+1} = g_{k+1}$ and let

$$\eta_3 = \lambda_k\kappa_{k+1} - \kappa_k\lambda_{k+1}, \quad \eta_4 = \alpha_k\lambda_{k+1} - \beta_k\kappa_{k+1}$$

From (7) (in theorem 2.16) we see that (η_3,η_4) is W_k allowable and it is easily checked that $\eta_{k,1}\eta_4 - \eta_{k,2}\eta_3 = g_k g_{k+1}$. Thus $\eta_{k,3},\eta_{k,4}$ do indeed exist.

Now suppose that $\eta_{k,3},\eta_{k,4}$ is any pair of integers satisfying the conditions stated in the theorem. Let $B = A_k S_k$ and note that $B \in M(w)$. We have

$$B = (\eta_{k,3}A_k + \eta_{k,4}B_k)/g_k$$

$$= ((\eta_{k,3}\alpha_k + \eta_{k,4}\kappa_k)(1,1) + (\eta_{k,3}\beta_k + \eta_{k,4}\lambda_k)W_k)/g_k$$

Let $\kappa = (\eta_{k,3}\alpha_k + \eta_{k,4}\kappa_k)/g_k$, $\lambda = (\eta_{k,3}\beta_k + \eta_{k,4}\lambda_k)/g_k$. Since $B \in M(w)$ we have $\kappa,\lambda \in Z(\delta)$ and it is easily checked that $\alpha_{k+1}\lambda - \beta_{k+1}\kappa = g_{k+1}$. We can therefore take $\kappa_{k+1} = \kappa$, $\lambda_{k+1} = \lambda$, $B_{k+1} = B$ and so we have

$$W_{k+1} = B_{k+1}/A_{k+1} = (B_{k+1}/A_k)/(A_{k+1}/A_k) = S_k/R_k$$

From the expressions for R_k, S_k we have

$$(1,1) = (\eta_{k,4}R_k - \eta_{k,2}S_k)/g_{k+1}, W_k = (-\eta_{k,3}R_k + \eta_{k,1}S_k)/g_{k+1}$$

Thus

$$M_k(w) = \{(\alpha(1,1) + \beta W_k)/g_k : (\alpha,\beta)\ W_k \text{ allowable}\}$$

$$= \{((\alpha\eta_{k,4} - \beta\eta_{k,3})R_k + (-\alpha\eta_{k,2} + \beta\eta_{k,1})S_k)/g_k g_{k+1}$$

$$: (\alpha,\beta)\ W_k \text{ allowable}\}$$

and so

$$M_{k+1}(w) = M_k(w)/R_k$$

$$= \{((\alpha\eta_{k,4} - \beta\eta_{k,3})(1,1) + (-\alpha\eta_{k,2} + \beta\eta_{k,1})W_{k+1})/g_k g_{k+1}$$

$$: (\alpha,\beta)\ W_k \text{ allowable}\}$$

$$= \{(\gamma(1,1) + \theta W_{k+1})/g_{k+1} : (\gamma,\theta)\ W_{k+1} \text{ allowable}\}$$

We must therefore have

$$I_{k+1} = \{(-\alpha\eta_{k,2} + \beta\eta_{k,1})/g_k : (\alpha,\beta)\ W_k \text{ allowable}\}$$

Now $\alpha = m_1 g_k + m_2\sigma_k' + m_4\psi_k$, $\beta = m_3 g_k + m_4\sigma_k$, $m_j \in Z$. (See theorem 2.16). Thus

$$I_{k+1} = \{-m_1\eta_{k,2} - m_2\sigma_k'\eta_{k,2}/g_k + m_3\eta_{k,1} + m_4(\sigma_k\eta_{k,1} - \psi_k\eta_{k,2})/g_k$$

$$: m_j \in Z\}$$

which is the required form for I_{k+1}.

Finally let (α,β) be W_k allowable and satisfy

$$(-\alpha\eta_{k,2} + \beta\eta_{k,1})/g_k = \sigma_{k+1}$$

We therefore have $((\alpha\eta_{k,4} - \beta\eta_{k,3})/g_k, \sigma_{k+1})\ W_{k+1}$ allowable. Now $(\psi_{k+1},\sigma_{k+1})$ is also W_{k+1} allowable and so

$$(\alpha\eta_{k,4} - \beta\eta_{k,3})/g_k - \psi_{k+1} \in I_{k+1}'$$

which is the final result of the theorem. //

We therefore calculate the required $M_{k+1}(w)$ representation as follows. First reduce I_{k+1} as given in theorem 2.19 to the standard form $Z[g_{k+1},\sigma_{k+1}]$. (See example 1.1). This will give $m_j, n_j \in Z$ such that

$$\begin{aligned} g_{k+1} &= m_1(-\eta_{k,2}) + m_2(-\sigma_k'\eta_{k,2}/g_k) + m_3\eta_{k,1} + m_4(\sigma_k\eta_{k,1} - \psi_k\eta_{k,2})/g_k \\ &= (\eta_{k,1}(m_3g_k + m_4\sigma_k) - \eta_{k,2}(m_1g_k + m_2\sigma_k' + m_4\psi_k))/g_k \\ \sigma_{k+1} &= (\eta_{k,1}(n_3g_k + n_4\sigma_k) - \eta_{k,2}(n_1g_k + n_2\sigma_k' + n_4\psi_k))/g_k \end{aligned} \qquad (11)$$

We can therefore take $\eta_{k,3} = m_1g_k + m_2\sigma_k' + m_4\psi_k$, $\eta_{k,4} = m_3g_k + m_4\sigma_k$. Clearly $(\eta_{k,3},\eta_{k,4})$ is W_k allowable. These coefficients are then used to calculate S_k and W_{k+1}. Now if $g_{k+1} = 1$ we know that $\psi_{k+1} = 0$ and so the complete representation of $M_{k+1}(w)$ is known. If $g_{k+1} > 1$ then we must determine ψ_{k+1}. From (11) we see that $\alpha = n_1g_k + n_2\sigma_k' + n_4\psi_k$, $\beta = n_3g_k + n_4\sigma_k$ satisfy the conditions in the theorem. We therefore calculate $\psi = (\alpha\eta_{k,4} - \beta\eta_{k,3})/g_k$ and then reduce this integer mod I_{k+1}' to obtain the required ψ_{k+1}. This completes the calculation of a standard representation of $M_{k+1}(w)$.

Note that when $\eta_{k,2} = 1$ we can bypass virtually all of the above calculation. It is not difficult to see that $\eta_{k,2} = 1$ can only occur if $g_k = g_{k+1} = 1$. Consequently we can take $\eta_{k,3} = -1$, $\eta_{k,4} = 0$. Furthermore we will also have $I_{k+1} = Z[1,\omega] = Z(\delta)$ and so $\sigma_{k+1} = \omega$, $\psi_{k+1} = 0$. Thus the only calculation required is that of $W_{k+1} = -(1,1)/R_k$. This short cut for finding a standard representation of $M_{k+1}(w)$ is particularly relevant when $Z(\delta)$ is a Euclidean domain ($d = -1, -2, -3, -7, -11$) since in almost all cases we find $\eta_{k,2} = 1$. For other small values of d we also find that the occurrence of $\eta_{k,2} = 1$ is relatively common.

Before presenting the above ideas in algorithmic form we consider one further point. It is quite possible that W_k will grow with

k eventually reaching an unmanageable size. The following theorem indicates how this possible problem can be avoided.

THEOREM 2.20

Suppose we have a standard representation

$$M_k(w) = \{(\alpha(1,1) + \beta W_k)/g_k : (\alpha,\beta)\ W_k \text{ allowable}\}$$

Then

$$M_k(w) = \{(\kappa(1,1) + \lambda W_k^{(1)})/g_k : (\kappa,\lambda)\ W_k^{(1)} \text{ allowable}\}$$

is a standard representation

if and only if

$$\exists\ \tau \in I_k' \text{ such that } W_k^{(1)} = (\tau/g_k)(1,1) + W_k$$

$$\text{and } \psi_k^{(1)} \equiv \psi_k - \tau\sigma_k/g_k \pmod{I_k'}$$

PROOF

We have $A_k W_k = B_k = \kappa_k(1,1) + \lambda_k W$ with $\alpha_k\lambda_k - \beta_k\kappa_k = g_k$. Thus

$$W = (\alpha_k B_k - \kappa_k A_k)/g_k$$

Now if we also have a standard representation

$$M_k(w) = \{(\kappa(1,1) + \lambda W_k^{(1)})/g_k : (\kappa,\lambda)\ W_k^{(1)} \text{ allowable}\}$$

then $A_k W_k^{(1)} = B_k^{(1)} = \kappa_k^{(1)}(1,1) + \lambda_k^{(1)} W$ with $\alpha_k\lambda_k^{(1)} - \beta_k\kappa_k^{(1)} = g_k$ and so

$$W = (\alpha_k B_k^{(1)} - \kappa_k^{(1)} A_k)/g_k$$

Equating these two expressions for W gives the relationship

$$B_k^{(1)} = ((\kappa_k^{(1)} - \kappa_k)/\alpha_k)A_k + B_k$$

which implies

$$W_k^{(1)} = ((\kappa_k^{(1)} - \kappa_k)/\alpha_k)\ (1,1) + W_k$$

Clearly $((\kappa_k^{(1)} - \kappa_k)/\alpha_k)(1,1) \in M_k(w)$ and so $\exists\ \tau \in I_k'$ such that $(\kappa_k^{(1)} - \kappa_k)/\alpha_k = \tau/g_k$. Note that $R = (\psi_k(1,1) + \sigma_k W_k)/g_k$, $R^{(1)} = (\psi_k^{(1)}(1,1) + \sigma_k W_k^{(1)})/g_k \in M_k(w)$ implies

$$R - R^{(1)} = (\psi_k - \psi_k^{(1)} - \tau\sigma_k/g_k)(1,1)/g_k \in M_k(w)$$

and so $\psi_k - \psi_k^{(1)} - \tau\sigma_k/g_k \in I_k'$ which is the second of the two conditions in the theorem.

The proof of the reverse implication is basically just a reversal of the steps in the previous paragraph. //

Thus to prevent the possibility of W_k growing with k we shall in practice always reduce the W_{k+1} obtained via theorem 2.19 as follows. Choose $\tau \in I_{k+1}'$ such that $|W_{k+1} + W_{k+1}^* + (2\tau/g_{k+1})(1,1)|$ is "small" and replace W_{k+1} with $W_{k+1} + (\tau/g_{k+1})(1,1)$. To see more clearly how τ should be chosen let $\tau = p_1 g_{k+1} + p_2 \sigma_{k+1}'$ and let $y_{k+1} = w_{k+1,1} + w_{k+1,2}$. We have

$$|W_{k+1} + W_{k+1}^* + (2\tau/g_{k+1})(1,1)| = 0$$

$$\Leftrightarrow \quad |y_{k+1} + 2\tau/g_{k+1}| = 0$$

$$\Leftrightarrow \quad \mathrm{Im}(y_{k+1} + 2\tau/g_{k+1}) = 0,\ \mathrm{Re}(y_{k+1} + 2\tau/g_{k+1}) = 0$$

$$\Leftrightarrow \quad \mathrm{Im}(y_{k+1}) - 2p_2|\delta|/c\, g_{k+1} = 0,\ \mathrm{Re}(y_{k+1}) + 2(p_1 + p_2 a_{k+1}/c\, g_{k+1}) = 0$$

$$\Leftrightarrow \quad p_2 = c\, g_{k+1}\, \mathrm{Im}(y_{k+1})/2|\delta|,\ p_1 = -\mathrm{Re}(y_{k+1})/2 - p_2 a_{k+1}/c\, g_{k+1}$$

We shall therefore choose τ by taking

$$p_2 = \{c\, g_{k+1}\, \mathrm{Im}(y_{k+1})/2|\delta|\},\ p_1 = \{-\mathrm{Re}(y_{k+1})/2 - p_2 a_{k+1}/c g_{k+1}\}$$

where $\{\ \}$ denotes the nearest integer function. (To choose τ so that "small" $\equiv$ minimal involves extra calculation when $a_{k+1} > 0$ and it is not justified. However when $a_{k+1} = 0$ the resulting W_{k+1} will have

$|W_{k+1} + W^*_{k+1}|$ minimal). Note that this process produces W_{k+1} satisfying $|W_{k+1} + W^*_{k+1}| < 2(g_{k+1} + |\delta|/c)/g_{k+1}$. Since $|W_{k+1} - W^*_{k+1}|$ does not vary with the choice of W_{k+1} (see the proof of theorem 2.18) we also have that $|W_{k+1}|, |W^*_{k+1}|$ will be approximately minimal amongst all possible choices for W_{k+1}.

The above reduction of W_{k+1} will be performed immediately following the calculation of W_{k+1} as outlined in the paragraph following theorem 2.19. Note that this reduction of W_{k+1} necessitates a modification of the calculation of ψ_{k+1}. We now calculate

$$\psi = (\alpha\eta_{k,4} - \beta\eta_{k,3})/g_k - \tau\sigma_{k+1}/g_{k+1}$$

and then reduce this integer mod I'_{k+1} to obtain the ψ_{k+1} corresponding to the reduced W_{k+1}.

ALGORITHM 2.1B

Given a standard representation of $M_k(w)$ and $R_k \in M_k(w)$ with $R_k = (\eta_{k,1}(1,1) + \eta_{k,2}W_k)/g_k$ we calculate a standard representation of $M_{k+1}(w)$ $(A_{k+1} = R_k A_k)$ as follows.

1 If $\eta_{k,2} = 1$ then set $\eta_{k,3} = -1$, $\eta_{k,4} = 0$, $g_{k+1} = 1$, $\sigma_{k+1} = \omega$ and go to 19

2 Set $\eta_1 = -\eta_{k,2}, \eta_2 = -\sigma'_k\eta_{k,2}/g_k$, $\eta_3 = \eta_{k,1}$, $\eta_4 = (\sigma_k\eta_{k,1} - \psi_k\eta_{k,2})/g_k$

3 Set $c_1 = (1,0,0,0)$, $c_2 = (0,1,0,0)$, $c_3 = (0,0,1,0)$, $c_4 = (0,0,0,1)$

4 Find ℓ such that $|\mathrm{Im}(\eta_\ell)| > 0$, $|\mathrm{Im}(\eta_\ell)|$ minimal

5 Exchange η_1, η_ℓ and c_1, c_ℓ

6 For $\ell = 2,3,4$ set $m = \{\mathrm{Im}(\eta_\ell)/\mathrm{Im}(\eta_1)\}$, replace

$$\eta_\ell \text{ with } \eta_\ell - m\eta_1, \quad c_\ell \text{ with } c_\ell - mc_1$$

7 If $|\mathrm{Im}(\eta_1)| > |\delta|/c$ go to 4

8 If $|\mathrm{Im}(\eta_1)| < 0$ then multiply η_1, c_1 by -1

9 Find ℓ, $2 \le \ell \le 4$ such that $\eta_\ell \neq 0$, $|\eta_\ell|$ minimal

10 Exchange η_2, η_ℓ and c_2, c_ℓ

11 For $\ell = 3,4$ set $m = \{\eta_\ell/\eta_2\}$, replace

$$\eta_\ell \text{ with } \eta_\ell - m\eta_2, \quad c_\ell \text{ with } c_\ell - mc_2$$

12 If $\eta_3 \neq 0$ or $\eta_4 \neq 0$ go to 9

13 If $\eta_2 < 0$ then multiply η_2, c_2 by -1

14 Set $g_{k+1} = \eta_2$

15 Set $m = [\mathrm{Re}(\eta_1)/g_{k+1}]$

16 Replace η_1 with $\eta_1 - mg_{k+1}$, c_1 with $c_1 - mc_2$

17 Set $\sigma_{k+1} = \eta_1$

18 Set $\eta_{k,3} = m_1 g_k + m_2 \sigma'_k + m_4 \psi_k$, $\eta_{k,4} = m_3 g_k + m_4 \sigma_k$

where $c_2 = (m_1, m_2, m_3, m_4)$

19 Set $S_k = (\eta_{k,3}(1,1) + \eta_{k,4} W_k)/g_k$, $W_{k+1} = S_k/R_k$

20 Set $p_2 = \{c\, g_{k+1}\, \mathrm{Im}(w_{k+1,1} + w_{k+1,2})/2|\delta|\}$

21 Set $p_1 = \{-\mathrm{Re}(w_{k+1,1} + w_{k+1,2})/2 - p_2 a_{k+1}/c\, g_{k+1}\}$

22 Replace W_{k+1} with $W_{k+1} + ((p_1 g_{k+1} + p_2 \sigma'_{k+1})/g_{k+1})(1,1)$

23 If $g_{k+1} > 1$ go to 25

24 Set $\psi_{k+1} = 0$ and stop

25 Set $\alpha = n_1 g_k + n_2 \sigma'_k + n_4 \psi_k$, $\beta = n_3 g_k + n_4 \sigma_k$

where $c_1 = (n_1, n_2, n_3, n_4)$

26 Set $\psi_{k+1} = (\alpha \eta_{k,4} - \beta \eta_{k,3})/g_k - (p_1 g_{k+1} + p_2 \sigma'_{k+1})\sigma_{k+1}/g_{k+1}$

27 Set $m = c\, \mathrm{Im}(\psi_{k+1})/|\delta|$

28 Replace ψ_{k+1} with $\psi_{k+1} + m\sigma'_{k+1}$

29 Set $m = [\psi_{k+1}/g_{k+1}]$

30 Replace ψ_{k+1} with $\psi_{k+1} - mg_{k+1}$

31 Stop //

We have now developed the major sections of our algorithm for calculating relative minima. Before stating the complete algorithm we consider one final point. The calculations required by the algorithm will mainly involve general complex arithmetic rather than $Z(\delta)$ integer arithmetic. Consequently we will obtain each A_k expressed as a two dimensional complex vector $A_k = (A_{k,1}, A_{k,2})$, $A_{k,j} \in \mathbb{C}$. However we are often interested in knowing the coefficients of A_k. The next theorem shows how α_k, β_k are most easily calculated from A_k.

THEOREM 2.21

Let A_k, $k > 0$ be a relative minimum of $M(w)$ with

$$A_k = (A_{k,1}, A_{k,2}),\ A_{k,j} \in \mathbb{C}$$

Then the coefficients of A_k are

$$\alpha_k = \{A_{k,1}/2\},\ \beta_k = \{A_{k,1}/2w\}$$

where $\{\ \}$ is the nearest integer function.

PROOF

We have $A_{k,1} = \alpha_k + \beta_k w$, $A_{k,2} = \alpha_k - \beta_k w$ and so

$$\alpha_k - A_{k,1}/2 = A_{k,2}/2,\ \beta_k - A_{k,1}/2w = -A_{k,2}/2w$$

Now $|A_{k,2}| = |A_k^*| < |A_0| = \min\{1, |w|\}$. Therefore $|A_{k,2}/2|, |A_{k,2}/2w| < 1/2$. The theorem now follows since $\alpha, \beta \in Z(\delta)$, $\alpha \neq \beta$ implies $|\alpha - \beta| \geq 1$. //

We now summarise the results of this section in the following algorithm.

ALGORITHM 2.1

Let $w \in \mathbb{C}\backslash\{0\}$. A half chain of relative minima of $M(w)$ can be

calculated as follows.

1 If $|w| \geq 1$ set $A_0 = (1,1)$, $\alpha_0 = 1$, $\beta_0 = 0$, $W_0 = W$, and go to 3

2 Set $A_0 = W$, $\alpha_0 = 0$, $\beta_0 = 1$, $W_0 = -W^{-1}$

3 Set $g_0 = 1$, $\sigma_0 = \omega$, $\psi_0 = 0$

4 Set $k = 0$

5 Use algorithm 2.1A to calculate $R_k \in M_k(w)$

6 Set $A_{k+1} = A_k R_k$, $\alpha_{k+1} = \{A_{k+1,1}/2\}$, $\beta_{k+1} = \{A_{k+1,1}/2w\}$

7 If $|R_k^*| = 0$ then stop

8 Use algorithm 2.1B to calculate a standard representation of $M_{k+1}(w)$

9 Increment k by 1

10 Go to 5 //

If $w \in Q(\delta)$ then this algorithm will eventually stop at step 7 having calculated a half chain of relative minima. Of course if $w \notin Q(\delta)$ then the algorithm is non-terminating since there are infinitely many A_k in a half chain for such a w. However in practice an appropriate stopping condition must be added to the algorithm. This is because calculations will necessarily involve finite approximations to irrational numbers and so we find a steady decline in the accuracy of the various values computed by the algorithm. In general we find that n significant digit arithmetic will only reliably calculate relative minima of magnitude up to approximately $10^{n/2}$ (for $|w| \approx 1$). Consequently to calculate large relative minima we must use multiprecision arithmetic. Note that it is only the calculation of $W_{k+1}, A_{k+1}, \alpha_{k+1}, \beta_{k+1}$ which require multiprecision arithmetic. We generally find that 6-8S is sufficient for all other calculations required by the algorithm, when $|\delta|$ is small.

We now illustrate the use of algorithm 2.1 with several examples. (We have used a fortran implementation of algorithm 2.1 on a Prime 750 computer for the examples in this thesis).

EXAMPLE 2.3

We repeat example 2.1. That is we have $w = \sqrt{1+\delta}$, $\delta = \sqrt{-10}$ and we calculate a chain $A_0, A_1, \ldots, A_7 \in M(w)$ using algorithm 2.1. The calculations in this example (and following examples) have been carried out using 14S arithmetic (double precision in fortran on the Prime 750) and rounded to the number of digits shown. Note that $d \equiv 2 \pmod 4$ and so $c = 1$, $\omega = \delta$.

We have

$$w \approx 1.469120 + 1.076249i$$

Since $|w| \geq 1$ we set

$$A_0 = (1,1),\ \alpha_0 = 1,\ \beta_0 = 0$$

and the representation of $M_0(w)$ is defined by

$$g_0 = 1,\ \psi_0 = 0,\ \sigma_0 = \delta,\ W_0 = (w,-w)$$

The calculation of $R_0 \in M(w)$ using algorithm 2.1A is summarised in the following table. (Note that initially $b(\beta) = 3.6|\delta|/c \approx 11.4$)

TABLE 2.3 Calculation of R_0

β	α	r^*	r	$b(\beta)$
1	-			
2	-			
3	$4+\delta$	.41	10.56	3.17
	$5+\delta$	.60	11.37	
δ				

This table and tables 2.4, 2.5, 2.6, 2.8, 2.9 give the following information. The first column gives all $\beta \in Z(\delta)^+ \cap I_k$ which required testing by algorithm 2.1A. A dash in the α column indicates that

$m_2 > m_2^+$ at step 5 and so no values of α needed to be considered. If $m_2 \le m_2^+$ then the α column lists all α corresponding to the pairings m_1, m_2 which needed checking. For each α we record the corresponding value of r^* and if $r^* < 1$ then r is also recorded. Finally if step 12 is reached then the updated value of $b(\beta)$ is recorded. Thus the coefficients of R_k correspond to the final updating of $b(\beta)$ in the table.

Thus $\eta_{0,1} = 4 + \delta$, $\eta_{0,2} = 3$ and so

$$R_0 = (4+\delta)(1,1) + 3W_0 \approx (8.41 + 6.39i, -.41 - .07i)$$

Therefore

$$A_1 = A_0 R_0 = R_0, \ \alpha_1 = 4+\delta, \ \beta_1 = 3$$

We now use algorithm 2.1B to calculate a representation of $M_1(w)$. We have

$$\eta_1 = -3, \ \eta_2 = 3\delta, \ \eta_3 = 4+\delta, \ \eta_4 = -10 + 4\delta$$

and so

$$I_1 = Z[\eta_1,\eta_2,\eta_3,\eta_4] = Z[1,\delta], \text{ that is } g_1 = 1, \ \sigma_1 = \delta$$

The reduction of I_1 to the standard representation also gives

$$c_2 = (-9, 0, -4, 1)$$

and so

$$\eta_{0,3} = (-9)g_0 + 0\,\sigma_0' + \psi_0 = -9, \ \eta_{0,4} = (-4)g_0 + \sigma_0 = -4 + \delta$$

$$S_0 = (-9)(1,1) + (-4+\delta)W_0 \approx (-18.28 + .34i, .28 - .34i)$$

and the initial value for W_1 is

$$W_1 = S_0/R_0 \approx (-1.36 + 1.07i, -.54 + .92i)$$

To reduce W_1 we calculate $p_2 = 0$, $p_1 = 1$ and reset W_1 to

$$(-1.36 + 1.07i, -.54 + .92i) + (1,1) = (-.36 + 1.07i, .46 + .92i)$$

Since $g_1 = 1$ we set $\psi_1 = 0$ and so the representation of $M_1(w)$ is defined by

$$g_1 = 1, \sigma_1 = \delta, \psi_1 = 0, W_1 \approx (-.36 + 1.07i, .46 + .92i)$$

We now increment k to 1 and use algorithm 2.1A to calculate $R_1 \in M_1(w)$. The results are summarised in the following table.

TABLE 2.4 Calculation of R_1

β	α	r^*	r	$b(\beta)$
1	-1	1.07		
	0	1.03		
2	-			
3	$-2-\delta$	.72	3.08	4.88
	$-1-\delta$	.55	2.08	3.68
δ	-			
$1+\delta$	$2-\delta$	.90	3.67	
	$3-\delta$	.94	3.31	
$1-\delta$	-4	.82	2.41	
	-3	.67	2.21	

Thus $\eta_{1,1} = -1-\delta$, $\eta_{1,2} = 3$, and we have

$$R_1 \approx (-2.08 + .06i, .39 - .39i)$$

$$A_2 = A_1R_1 \approx (-17.81 - 12.78i, -.19 + .13i)$$

$$\alpha_2 = -9 - 2\delta, \beta_2 = -6$$

The calculation of a representation of $M_2(w)$ is once again straight forward and we obtain

$$I_2 = Z[-3,3\delta,-1-\delta,10-\delta] = Z[1,\delta], c_2 = (-4,0,1,-1)$$

$$\eta_{1,3} = -4, \eta_{1,4} = 1 - \delta$$

$$S_1 = -4(1,1) + (1-\delta)W_1 \approx (-.96 + 2.21i, -.61 - .54i)$$

$$W_2 = S_1/R_1 \approx (.49 - 1.05i, -.09 - 1.48i)$$

We have $p_2 = p_1 = 0$ and so W_2 is left unchanged. Since $g_2 = 1$ we have $\psi_2 = 0$ and so the representation of $M_2(w)$ is defined by

$$g_2 = 1,\ \sigma_2 = \delta,\ \psi_2 = 0,\ W_2 = (.49 - 1.05i,\ -.09 - 1.48i)$$

We now increment k to 2 and calculate $R_2 \in M_2(w)$. The search for R_2 is summarised in the following table.

TABLE 2.5 Calculation of R_2

β	α	r^*	r	$b(\beta)$
1	-			
2	δ	.28	1.45	3.36
	$1+\delta$	.84	2.25	
3	-			
δ	-5	.44	2.29	
	-4	.74	1.70	
$1+\delta$	-			
$1-\delta$	-			

Thus $\eta_{2,1} = \delta$, $\eta_{2,2} = 2$ and we have

$$R_2 = \delta(1,1) + 2W_2 \approx (.99 + 1.06i,\ -.19 + .20i)$$

$$A_3 = A_2R_2 \approx (-4.01 - 31.56i,\ .0079 - .062i)$$

$$\alpha_3 = -2 - 5\delta,\ \beta_3 = -6 - 2\delta$$

The calculation of a representation of $M_3(w)$ is slightly more involved than previous cases since we find $g_3 = 2$. We have

$$I_3 = Z[-2,2\delta,\delta,-10] = Z[2,\delta],\ c_2 = (-1,0,0,0)$$

$$\eta_{2,3} = -1,\ \eta_{2,4} = 0$$

$$S_2 = -(1,1) + 0W_2 = (-1,-1)$$

$$W_3 = S_2/R_2 \approx (-.47 + .50i,\ 2.47 + 2.66i)$$

We then calculate $p_2 = 1$, $p_1 = -1$ and replace W_3 with

$$W_3 + ((-g_3+\sigma_3')/g_3)(1,1) \approx (-1.47 - 1.08i, 1.47 + 1.08i)$$

We also have $c_1 = (0,0,1,0)$ giving $\alpha = 0$, $\beta = 1$ and so

$$\psi_3 \equiv -\eta_{2,3}/g_2 - (-g_3+\sigma_3')\sigma_3/g_3 \pmod{I_3'}$$

$$\equiv 1 - (-2-\delta)\delta/2 \pmod{I_3'}$$

$$\equiv -4+\delta \equiv 0 \pmod{I_3'}$$

Therefore the required representation of $M_3(w)$ is defined by

$$g_3 = 2,\ \sigma_3 = \delta,\ \psi_3 = 0,\ W_3 \approx (-1.47 - 1.08i, 1.47 + 1.08i)$$

The calculation of $R_3 \in M_3(w)$ is quite short as is indicated by the following table.

TABLE 2.6 Calculation of R_3

β	α	r^*	r	$b(\beta)$
2	$-4-\delta$ $-2-\delta$	.73 .69	4.37 3.63	2.95 2.54

Thus $\eta_{3,1} = -2 - \delta$, $\eta_{3,2} = 2$ and we have

$$R_3 = ((-2-\delta)(1,1) + 2W_3)/2 \approx (-2.47 - 2.66i, .47 - .50i)$$

$$A_4 = A_3 R_3 \approx (-73.97 + 88.58\,i, -.028 - .033i)$$

$$\alpha_4 = -37 + 14\delta,\ \beta_4 = -2 + 10\delta$$

The remaining calculations for this example involve nothing new and so we briefly list the results.

Calculation of $M_4(w)$ representation

$$I_4 = Z[-2,\delta,-2-\delta,5-\delta] = Z[1,\delta],\ c_2 = (-4,0,1,-1)$$

$$\eta_{3,3} = -8,\ \eta_{3,4} = 2-\delta,\ S_3 \approx (-7.17 + 1.25i, -.83 - 1.25i)$$

W_4 (initially) $\approx (1.09 - 1.68i, .51 - 2.11i)$, $p_2 = -1$, $p_1 = -1$

and so $M_4(w)$ is defined by

$g_4 = 1$, $\sigma_4 = \delta$, $\psi_4 = 0$, $W_4 \approx (.09 + 1.48i, -.49 + 1.05i)$

Calculation of $R_4 \in M_4(w)$

$\beta = 1,2,3,\delta,1+\delta, 2+\delta, 1-\delta, 2-\delta$ required testing

$\eta_{4,1} = 1-\delta$, $\eta_{4,2} = 3$, $R_4 \approx (1.28+1.28i,-.48-.01i)$

$A_5 = A_4R_4 \approx (-208.01+18.96i,.013+.016i)$

$\alpha_5 = -104 + 3\delta$, $\beta_5 = -43 + 12\delta$

Calculation of $M_5(w)$ representation

$I_5 = Z[-3,3\delta,1-\delta,10+\delta] = Z[1,\delta]$, $c_2 = (-4,0,-1,-1)$

$\eta_{4,3} = -4$, $\eta_{4,4} = -1-\delta$, $S_4 \approx (.59 - 1.78i,-.19+.51i)$

$W_5 \approx (-.46 - .92i,.36 - 1.07i)$, $p_2 = p_1 = 0$

and so $M_5(w)$ is defined by

$g_5 = 1$, $\sigma_5 = \delta$, $\psi_5 = 0, W_5 \approx (-.46 - .92i, .36 - 1.07i)$

Calculation of $R_5 \in M_5(w)$

$\beta = 1,2,3,4,\delta,1+\delta,2+\delta,1-\delta,2-\delta$ required testing

$\eta_{5,1} = -1+\delta$, $\eta_{5,2} = 3$, $R_5 \approx (-2.39+.39i,.075-.057i)$

$A_6 = A_5R_5 \approx (489.998 - 126.492i, .0019+.00049i)$

$\alpha_6 = 245-20\delta$, $\beta_6 = 88-34\delta$

Calculation of $M_6(w)$ representation

$I_6 = Z[-3,3\delta,-1+\delta,-10-\delta] = Z[1,\delta]$, $c_2 = (-4,0,1,1)$

$\eta_{5,3} = -4$, $\eta_{5,4} = 1+\delta$, $S_5 \approx (-1.54-2.39i, -.25+.06i)$

W_6 (initially) $\approx (.47+1.08i, -2.47-1.08i)$, $p_2 = 0$, $p_1 = 1$

and so $M_6(w)$ is defined by

$$g_6 = 1,\ \sigma_6 = \delta,\ \psi_6 = 0,\ W_6 \approx (1.47+1.08i,-1.47-1.08i)$$

Calculation of $R_6 \in M_6(w)$

$\beta = 1,2,3,\delta$ required testing

$$\eta_{6,1} = 4+\delta,\ \eta_{6,2} = 3,\ R_6 \approx (8.41+6.39i,-.41-.07i)$$

$$A_7 = A_6R_6 \approx (4928.0007+2068.13i,-.00075-.00033i)$$

$$\alpha_7 = 2464 + 327\delta,\ \beta_7 = 1427 - 108\delta$$

We now note several points about these calculations. The first point to note is the amount of work involved in calculating the chain $A_0, A_1, \ldots, A_7$. In example 2.1 we had to test approximately 10^6 possibilities. However in using algorithm 2.1 we only had to consider 38 values of β. Of course the calculation of each $M_k(w)$ also involves a reasonable amount of work but it is clear that algorithm 2.1 is a vast improvement over the exhaustive search technique used in example 2.1.

Note that the chains in example 2.1 and this example are not identical. However the only difference is a root of unity factor and therefore of no real consequence.

Another point to note is that we appear to have $W_6 = W_0$ and $R_6 = R_0$. Since we have used finite precision arithmetic in these calculations we must be cautious about drawing any conclusions from such observations. However if we recall from example 2.1 that A_6 is a unit of M(w) (which is closed under multiplication) it is not difficult to see that we will in fact have

$$M_6(w) = M_0(w) = M(w)$$

It is then easily seen that $W_6 = W_0$, $R_6 = R_0$.

Finally we note that we also appear to have

$$W_1 = -W_5^*,\ W_2 = -W_4^*,\ W_3 = -W_3^*$$

Since $I_k = I_{6-k}$, $\psi_k = \psi_{6-k}$ for $k = 1,2,3$ this would be equivalent to having

$$M_1(w) = M_5^*(w),\ M_2(w) = M_4^*(w),\ M_3(w) = M_3^*(w)$$

where

$$M_k^*(w) = \{R^* : R \in M_k(w)\}$$

These observations are also easily confirmed algebraically and are typical of the case $w = \sqrt{\alpha}$, $\alpha \in Q(\delta)$. (See section four of this chapter). //

The value of w in the next example has been carefully chosen to illustrate a number of points concerning relative minima of M(w). It also gives another illustration of the use of algorithm 2.1.

EXAMPLE 2.4

Let $\delta = \sqrt{-17}$, $w = .72 + .21\delta \approx .72 + .8658521814i$. Algorithm 2.1 produces the results listed in table 2.7. (Note that $-17 \equiv 3 \pmod 4$ and so $c = 1$, $\omega = \delta$. Furthermore $|w| \geq 1$ so $A_0 = (1,1)$, $\alpha_0 = 1$, $\beta_0 = 0$)

TABLE 2.7 Calculation of relative minima of M(w)

	$M_k(w)$			R_k		A_{k+1}		
k	g_k	ψ_k	σ_k	$\eta_{k,1}$	$\eta_{k,2}$	$\lvert A_{k+1}\rvert$	α_{k+1}	β_{k+1}
0	1	0	δ	1	1	1.9	1	1
1	1	0	δ	$1+\delta$	3	9.3	$-2+\delta$	$1+\delta$
2	3	1	$1+\delta$	$1-\delta$	6	9.6	$3+\delta$	4
3	2	0	$1+\delta$	4	$3-\delta$	9.7	$2-\delta$	$-2-\delta$
4	1	0	δ	$-1+\delta$	4	9.9	$3-\delta$	$-1-\delta$
5	2	1	$1+\delta$	-2	2	16.5	$1+2\delta$	$6+\delta$
6	1	0	δ	$1+\delta$	3	24.2	12	$7-2\delta$
7	3	0	$1+\delta$	3	$2-\delta$	29.4	$-8+3\delta$	$4+3\delta$
8	1	0	δ	$-1+\delta$	3	30.8	$-13-2\delta$	$-13+\delta$
9	3	0	$2+\delta$	$1+\delta$	3	82.5	$2-10\delta$	$-27-6\delta$
10	3	2	$1+\delta$	$1-\delta$	3	111.6	$-45+8\delta$	$-3+12\delta$
11	1	0	δ	$-\delta$	3	175.7	$-47+18\delta$	$24+18\delta$
12	1	0	δ	δ	2	225.2	$-72-21\delta$	-100

The algorithm stops at $k = 12$ since $|R_{12}^*| = |A_{13}^*| = 0$. This is as we

would expect in view of theorem 2.9 since $w \in Q(\delta)$. Note that $w = \alpha_{13}/\beta_{13}$. In comparison with example 2.3 we find in the present example that there is a greater proportion of non-basic relative minima and that $|A_k|$ grows more slowly with k. To give some idea of the amount of work involved in the calculation of the various R_k tables 2.8, 2.9 detail two extreme cases.

TABLE 2.8 Calculation of R_2

β	α	r^*	r	$b(\beta)$
3	$-2-\delta$	.75	1.34	8.68
	$1-\delta$	.50	1.22	8.24
6	$-2-\delta$	.91	1.15	7.99
	$1-\delta$	.75	1.03	7.54
	$-1-2\delta$	.78	2.36	
	$2-2\delta$	1.003		
$1+\delta$	4	.36	1.06	
	7	1.01		
$4+\delta$	$2-\delta$	.71	1.15	
	$5-\delta$	.29	1.77	
$2-\delta$	-7	1.14		
	-4	.66	1.05	
	$-6-\delta$	.93	2.12	
	$-3-\delta$	.82	1.44	
$5-\delta$	$-6-\delta$	.65	1.99	
	$-3-\delta$	.51	1.26	

TABLE 2.9 Calculation of R_9

β	α	r^*	r	$b(\beta)$
3	$1+\delta$	.56	2.68	4.09
	$4+\delta$	.62	3.30	

The amount of calculation required to find R_k clearly varies quite considerably. The remaining 11 cases are fairly evenly distributed between these two extremes. Note that for $\beta = 6$, $2 - \delta$ in table 2.8 we have in each case 4 possibilities for α to consider. This corresponds to $m_2^+ = m_2 + 1$ in step 4 of algorithm 2.1A, that is there are two possibilities for m_2.

We now note several points of a general nature which this example was chosen to illustrate. Firstly we note that

$$|\beta_3| = 4 < \sqrt{18} = |\beta_2|, \quad |\beta_5| = \sqrt{18} < \sqrt{21} = |\beta_4|$$

Thus although we generally have $|\beta_{k+1}| \geq |\beta_k|$ this is obviously not always the case. It is for this reason that theorem 2.13 was stated in terms of $|\beta| = \max\{|\beta_k|, |\beta_{k+1}|\}$.

The second point to note is that although the equation

$$(71+21\delta)\theta - 100\lambda = \alpha_{13}\theta - \beta_{13}\lambda = 1$$

is solvable ($\theta = 12-41\delta, \lambda = 155-27\delta$ is one solution) the coefficients of A_{12} do not give a solution. In fact $\alpha_{13}\beta_{12} - \alpha_{12}\beta_{13} = -2$. Thus given $\alpha,\beta \in Z(\delta)$ such that the equation $\alpha\theta - \beta\lambda = 1$ is solvable for $\theta,\lambda \in Z(\delta)$ we see that the relative minima of $M(\alpha/\beta)$ are not guaranteed to produce a solution of this equation. This contrasts with the simple continued fraction case where the penultimate convergent of the expansion of a/b, $a,b \in Z$, $(a,b) = 1$ is guaranteed to give a solution $x,y \in Z$ of the equation $ax - by = 1$. The reason why $M(\alpha/\beta)$ can fail to produce a solution of the equation $\alpha\theta - \beta\lambda = 1$ is as follows. Given a solvable equation of this type we can only guarantee that there is a solution with θ satisfying $|\theta| < t\beta$, $t = |1+\delta|/2$ for $d \equiv 2,3 \pmod 4$, $t = (1+|\delta|^2)/4|\delta|$ for $d \equiv 1 \pmod 4$. Now when $t > 1$ it is therefore possible that all solutions of the equation have $|\theta| > |\beta|$. Consequently it is possible that for all solutions of the equation we have

$$|B| > |A|, \quad B = \lambda(1,1) + \theta W, \quad A = \alpha(1,1) + \beta W$$

where $w = \alpha/\beta$. Since $|A^*| = 0$ it follows that in such cases B can not be a relative minimum of M(w). (Note that even if the relative minima of $M(\alpha/\beta)$ were guaranteed to give a solution of $\alpha\theta - \beta\lambda = 1$ then this would not provide the most efficient method for solving the equation. Equations of the form $\alpha\theta - \beta\lambda = \kappa$ are most easily solved using the ideas used in the first part of algorithm 2.1B).

The final point we note concerns the relationship between the relative minima of $M(w)$ and the relative minima of $M(w+\alpha)$, $\alpha \in Z(\delta)$. We use a fairly arbitrary choice of $\alpha = 9$ to illustrate the situation.

TABLE 2.10 Relative minima of $M(w+9)$

k	0	1	2	3	4	5
α_k^+	1	10	$39+\delta$	$-6-10\delta$	$55+11\delta$	$75-18\delta$
β_k^+	0	1	4	$-1-\delta$	$6+\delta$	$7-2\delta$
	6	7	8	9	10	11
	$28+30\delta$	$-130+7\delta$	$-241-64\delta$	$-72+116\delta$	$169+180\delta$	$-972-21\delta$
	$4+3\delta$	$-13+\delta$	$-27-6\delta$	$-3+12\delta$	$24+18\delta$	-100

As we might expect there is a strong link between the relative minima of $M(w)$ and $M(w+9)$. It is easily checked that

$$\alpha_k^+ = \alpha_k + 9\beta_k,\ \beta_k^+ = \beta_k,\ k = 0,1$$

$$\alpha_2^+ = \alpha_3 + 9\beta_3,\ \beta_2^+ = \beta_3$$

$$\alpha_k^+ = \alpha_{k+2} + 9\beta_{k+2},\ \beta_k^+ = \beta_{k+2},\ k = 3,4,\ldots,11$$

Note that there are no relative minima of $M(w+9)$ corresponding to $A_2, A_4 \in M(w)$. This illustrates the fact that there is not necessarily a one to one correspondence between the relative minima of $M(w+\alpha)$ and $M(w)$ for $\alpha \in Z(\delta)$. This contrasts once again with the simple continued fraction case where we find a one to one correspondence between the convergents of $x \in \mathbf{R}$ and the convergents of $x+m$, $m \in Z$. Note that the β_k^+ coefficients in table 2.10 satisfy $|\beta_k^+| \leq |\beta_{k+1}^+|$ and in fact it can be shown that $|\beta_{k+1}| < |\beta_k|$ can only occur for $|w| \approx 1$. //

We finish this section by briefly considering the special case where $h(d) = 1$, that is $Z(\delta)$ is a unique factorization domain. This of course occurs for

$$d = -1,-2,-3,-7,-11,-19,-43,-67,-163$$

We have already noted that the relative minima of modules over such $Z(\delta)$ will necessarily be basic. It therefore follows that a standard representation of any $M_k(w)$ will always have $g_k = 1$, $\psi_k = 0$, $\sigma_k = \omega$. Consequently algorithm 2.1 can be modified to take advantage of this fact. Modifications include deletion of all redundant steps and arithmetic. For example steps 14 to 17 and 23 to 30 in algorithm 2.1B can be deleted and a step such as 10 in algorithm 2.1A can be rewritten as

10 Calculate $r = |m_1 + m_2\omega' + \beta w_{k,1}|$

The simplified algorithm which results will of course involve less work than the original algorithm. However the improvement will be relatively small since the main work in algorithm 2.1 lies in the calculation of R_k, $\eta_{k,3}$, $\eta_{k,4}$ which must of course still be calculated in the simplified version of the algorithm.

SECTION FOUR

PERIODIC RELATIVE MINIMA

In this section we shall mainly confine our attention to the case $w^2 \in Q(\delta)$, $w \notin Q(\delta)$. Results developed are in the main generalizations of results associated with periodic simple continued fractions. A specific algorithm for the calculation of periodic relative minima will be derived. (This algorithm plus a modified version of this algorithm will be used in the next chapter to calculate fundamental units of certain quartic fields of the form $Q(\delta)(\sqrt{\gamma})$, $\gamma \in Z(\delta)$). Finally we shall complete the development of the bound for $|A_k^*|$ which was stated and partially proved in section two of this chapter.

We begin with a theorem which leads to the definition of periodic relative minima.

THEOREM 2.22

Suppose there exist $A_m, A_n \in M(w)$ with $m \neq n$ such that $M_m(w) = M_n(w)$. Then the following results hold.

(a) For all $k \in Z$ we have $E_{m,m+k} = E_{n,n+k}$. Therefore given $R_{m,m+k} \in M_m(w)$ we can choose $R_{n,n+k} \in M_n(w)$ such that $R_{n,n+k} = R_{m,m+k}$. (Note however that given an $R_{m,m+k}$ and an $R_{n,n+k}$ we cannot assume that they are equal although we certainly must have $R_{m,m+k} \sim R_{n,n+k}$).

(b) We can choose A_q, $q \in Z^+$ such that $M_0(w) = M_q(w)$.

(c) M(w) has an infinite complete chain of relative minima such that $M_k(w) = M_{k+q}(w)$, $\forall\ k \in Z$.

PROOF

(a) The relative minima of $M_j(w)$ satisfy

$$\ldots < |R_{j,j-1}| < |R_{j,j}| = 1 < |R_{j,j+1}| < \ldots \qquad (12)$$

Now if $M_m(w) = M_n(w)$ then $M_m(w)$, $M_n(w)$ must have identical relative minima. It therefore follows from (12) that we must have $E_{m,m+k} = E_{n,n+k}$, $\forall\ k \in Z$.

(b) Assume $n > m$. Since A_m, A_0 exist so does $R_{m,0} = A_0/A_m$. We can therefore choose $R_{n,n-m}$ such that $R_{n,n-m} = R_{m,0}$. (Take $k = -m$ in part (a)). Set $A_{n-m} = A_n R_{n,n-m} = A_n R_{m,0}$ and set $q = n-m$. We therefore have $A_q = A_n R_{m,0}$ with $q \in Z^+$. Now

$$M_0(w) = M(w)/A_0 = M(w)/A_m R_{m,0} = M_m(w)/R_{m,0}$$

and

$$M_q(w) = M(w)/A_q = M(w)/A_n R_{m,0} = M_n(w)/R_{m,0}$$

Since $M_m(w) = M_n(w)$ it follows that $M_0(w) = M_q(w)$.

(c) Choose any $A_1, A_2, .., A_{q-1} \in M(w)$. We show that the required chain can be generated inductively from the finite chain $A_0, A_1, \ldots, A_q$. We begin by extending this initial chain into a half chain which satisfies $M_k(w) = M_{k+q}(w)$, $\forall\ k \geq 0$.

Suppose that the chain $A_0, A_1, \ldots, A_j, \ldots, A_{j+q}$ satisfies $M_k(w) = M_{k+q}(w)$ for $k = 0,1,\ldots,j$. Since A_j, A_{j+1} exist we see that $R_j = A_{j+1}/A_j$ exists. Now $M_j(w) = M_{j+q}(w)$ and so by part (a) we see that we can choose $R_{j+q} = R_j$. Set $A_{j+q+1} = A_{j+q}R_{j+q}$, that is $A_{j+q+1} = A_{j+q}R_j$. It is now easily shown that $M_{j+1}(w) = M_{j+q+1}(w)$. (See part (b) where we showed that $M_0(w) = M_q(w)$). The process outlined can be repeated indefinitely thereby producing the required infinite half chain.

This half chain can then be extended into the required infinite complete chain by using a similar process to define A_j, $j = -1, -2, \ldots$ (Briefly, for $j = 0, -1, -2, \ldots$ take $R_{j+q,j+q-1} = A_{j+q-1}/A_{j+q}$, choose $R_{j,j-1} = R_{j+q,j+q-1}$ and set $A_{j-1} = A_j R_{j,j-1} = A_j R_{j+q,j+q-1}$). //

In view of this theorem we make the following definition.

DEFINITION 2.11

We shall say that M(w) has *periodic relative minima* if and only if $M_0(w) = M_q(w)$ for some $q \in Z^+$.

A chain of relative minima $A_0, A_1, \ldots, A_q$, $q \in Z^+$ which satisfies $M_0(w) = M_q(w)$, $M_0(w) \neq M_k(w)$ for $0 < k < q$ will be called a *period* of relative minima. The integer q will be called the *length* of the period. //

Theorem 2.22 effectively shows that M(w) has periodic relative minima if and only if $M_m(w) = M_n(w)$ for some $m,n \in Z, m \neq n$. Thus the definition of periodic relative minima is not as restrictive as it may initially appear. Furthermore periodicity is clearly always pure periodicity. Note that the phrase 'periodic relative minima' is slightly misleading. It is of course the ratios $R_k = A_{k+1}/A_k$ and the modules $M_k(w)$ which are actually periodic. Finally note that if M(w) has periodic relative minima then the proof of theorem 2.22c shows that the calculation of a complete chain of relative minima is effectively

achieved by the calculation of a period of relative minima. (However see theorem 2.29 for a more straightforward way of generating a complete chain from a period). Consequently the modified version of algorithm 2.1 which we develop for periodic relative minima will confine its attention to the calculation of a period of relative minima.

We now illustrate some of the ideas presented so far in this section.

EXAMPLE 2.5

(a) Recall examples 2.1, 2.3. In example 2.3 we saw that for $w = \sqrt{1+\delta}$, $\delta = \sqrt{-10}$ we have $M_0(w) = M_6(w)$. Thus $M(w)$ has periodic relative minima. Furthermore it is not difficult to see that $M_0(w) \neq M_k(w)$, $0 < k < 6$. Thus $A_0, A_1, \ldots, A_6$ form a period of relative minima of length 6. In example 2.1 we noted that a complete chain of relative minima of $M(w)$ can be generated from this period. This illustrates the result of theorem 2.22c although the method of generation of a complete chain in example 2.1 differs from the method given in the proof of theorem 2.22c. (However the method of generation in example 2.1 is the simpler of the two and we shall generalise it in theorem 2.29).

(b) Let $\delta = \sqrt{-10}$, $w = \sqrt{(3+2\delta)/7}$. With the aid of algorithm 2.1 we find

$$A_0 = (1,1)$$

$$A_1 = (1,1) + W$$

$$A_2 = (-1+\delta)(1,1) + (1+\delta)W$$

$$A_3 = 6(1,1) + (5-\delta)W$$

and it is relatively easy to check that $M_3(w) = M_0(w)$. Thus A_0, A_1, A_2, A_3 is a period of relative minima of $M(w)$ of length 3. (Note that $Q(\delta)(w) = Q(\sqrt{-10})(\sqrt{35})$ and $A_3 = 6(1,1) + (\sqrt{35}, -\sqrt{35})$ corresponds to the real quadratic unit $6 + \sqrt{35}$).

Since $|w| = 1$ we have

$$A_3^{(1)} = WA_3 = (5+\delta)(1,1) + 6W$$

is also a relative minimum of M(w). However

$$(7/(3+2\delta))W = W^{-1} = A_3/A_3^{(1)} \in M_3^{(1)}(w)$$

and it follows that $M_3^{(1)}(w) \neq M_0(w)$. Consequently it is not difficult to check that

$$A_0, A_1, A_2, A_3^{(1)}, A_4 = A_3A_1, A_5 = A_3A_2, A_6 = (A_3)^2$$

is a period of relative minima of length 6. //

Example 2.5b illustrates the fact that the integer q in definition 2.11 depends on the period under consideration. However it is not difficult to deduce using the ideas in the proof of theorem 2.22b that for a given M(w) there exists a period (not unique) of minimal length and that the length of any other period is a multiple of this minimal length. In view of these facts it is clear from an efficiency point of view that the modified version of algorithm 2.1 which we will develop later in this section should be designed to calculate a period of minimal length. Consequently we will in future mainly confine our attention to such minimal periods.

We now develop conditions which w must satisfy if M(w) is to have periodic relative minima.

THEOREM 2.23

If M(w) has periodic relative minima then $w^2 \in Q(\delta)$, $w \notin Q(\delta)$. Thus w can be written as

$$w = \sqrt{\gamma/h},\ \gamma \in Z(\delta),\ \ h \in Z^+$$

with γ/h not a perfect square in $Q(\delta)$.

PROOF

Suppose M(w) has periodic relative minima. Then $\exists\ q \in Z^+$ such that $M_0(w) = M_q(w)$. Therefore we can choose $R_0 \in M_0(w)$ and $R_q \in M_q(w)$ such that $R_0 = A_1/A_0 = A_{q+1}/A_q = R_q$, or equivalently $A_1 A_q = A_0 A_{q+1}$. If we now substitute $A_k = (\alpha_k + \beta_k w, \alpha_k - \beta_k w)$, $k = 0, 1, q, q+1$ in this equation and

multiply out then we obtain two equalities. (One from each component).

$$\alpha_1\alpha_q + \beta_1\beta_q w^2 + (\alpha_1\beta_q + \alpha_q\beta_1)w = \alpha_0\alpha_{q+1} + \beta_0\beta_{q+1}w^2 + (\alpha_0\beta_{q+1} + \alpha_{q+1}\beta_0)w$$

$$\alpha_1\alpha_q + \beta_1\beta_q w^2 - (\alpha_1\beta_q + \alpha_q\beta_1)w = \alpha_0\alpha_{q+1} + \beta_0\beta_{q+1}w^2 - (\alpha_0\beta_{q+1} + \alpha_{q+1}\beta_0)w$$

When we add these two equalities and rearrange we obtain

$$(\beta_1\beta_q - \beta_0\beta_{q+1})w^2 + (\alpha_1\alpha_q - \alpha_0\alpha_{q+1}) = 0 \qquad (13)$$

Now $\alpha_0\beta_0 = 0$. Therefore if $\beta_1\beta_q - \beta_0\beta_{q+1} = \alpha_1\alpha_q - \alpha_0\alpha_{q+1} = 0$ then either $\alpha_1\alpha_q = 0$ or $\beta_1\beta_q = 0$. However it is easily seen that $\alpha_k, \beta_k \neq 0$ for $k \neq 0$. Thus $\beta_1\beta_q - \beta_0\beta_{q+1}, \alpha_1\alpha_q - \alpha_0\alpha_{q+1} \neq 0$ and so (13) implies that $w^2 \in Q(\delta)$. Theorem 2.22c plus theorem 2.9 together show that $w \notin Q(\delta)$ and the theorem now follows. //

This theorem shows that our definition of periodicity is more restrictive than simple continued fraction periodicity. Recall that simple continued fraction periodicity is effectively defined in terms of the ratios $x_{k+1} = -(p_{k-1}-q_{k-1}x)/(p_k-q_kx)$ eventually becoming periodic. (This would generalize to M(w) by defining periodicity in terms of the second components of the R_k eventually becoming periodic). The ratios $(p_k+q_kx)/(p_{k-1}+q_{k-1}x)$ become periodic only if $x^2 \in Q$, $x \notin Q$, but since the simple continued fraction algorithm makes no use of these ratios there is no distinction made between the two types of periodicity. However algorithm 2.1 makes use of both components of $R \in M_k(w)$ and consequently the two types of periodicity are clearly distinguished with the more restrictive type being easier to work with. (As well as being more natural in the present context).

In the rest of this section we shall mainly be dealing with w of the form $\sqrt{\gamma/h}$ with γ,h as in theorem 2.23. Therefore in the rest of this section when we write $\sqrt{\gamma/h}$ it will be assumed that γ,h satisfy the conditions given in theorem 2.23.

Our purpose in the next few theorems is to prove the converse of theorem 2.3, that is to prove that $M(\sqrt{\gamma/h})$ has periodic relative minima. We begin by noting that if $w = \sqrt{\gamma/h}$ then the components of $A = (\alpha+\beta w, \alpha-\beta w) \in M(w)$ are conjugates in the field $Q(\delta)(w)$ which is a quadratic extension of $Q(\delta)$. This result also extends to $R \in M(w)/B$ for $B \in M(w)$. We therefore make the following definition.

DEFINITION 2.12

Let $w = \sqrt{\gamma/h}$ and $B \in M(w)$, $|BB^*| \neq 0$. We define $N_\delta(R)$ for $R \in M(w)/B$ to be the first component of RR^*. //

Of course both components of RR^* are identical and so we have $N_\delta(R)(1,1) = RR^*$. The following results are obvious.

THEOREM 2.24

Let $w = \sqrt{\gamma/h}$. Then N_δ corresponds to the relative norm function from $Q(\delta)(w)$ to $Q(\delta)$. In particular for $A = (\alpha+\beta w, \alpha-\beta w)$, $B \in M(w)$ we have

(a) $N_\delta(A) = \alpha^2 - \beta^2\gamma/h \in Q(\delta), hN_\delta(A) \in Z(\delta)$

(b) $N_\delta(A) = N_\delta(A^*)$

(c) $N_\delta(AB) = N_\delta(A)\, N_\delta(B)$

(d) $N_\delta(A/B) = N_\delta(A)/N_\delta(B)$, provided $|BB^*| \neq 0$ //

THEOREM 2.25

Let A_k, $k > 0$ be a relative minimum of $M(w)$, $w = \sqrt{\gamma/h}$. We have

$$|N_\delta(A_k)| < \frac{4\sqrt{2}|\delta||w|}{c\sqrt[4]{3}\,\pi^{\frac{1}{2}}} + \frac{r}{|\beta_k|}$$

where $r \in \mathbf{R}^+$ depends only on $|\delta|$ and $|w|$.

PROOF

From theorem 2.13 we have

$$|A_k^*| < \frac{2\sqrt{2}|\delta|}{c\sqrt{3}\pi^{\frac{1}{2}}|\beta_k|} + \frac{9|\delta|^2}{|\beta_k|^2}$$

and from theorem 2.1 we have

$$|A_k| = |\alpha_k+\beta_k w| \leq 2|\beta_k w| + |A_k^*|$$

Since $|N_\delta(A_k)| = |A_k||A_k^*|$ the result now follows easily. //

The representation of W_k given in the next theorem not only enables us to prove the required converse of theorem 2.23 but will also be of considerable importance in the development (later in this section) of the modified version of algorithm 2.1. It will enable us to reduce the need for multiprecision arithmetic when $|A_k|$ is large and it will also enable us to recognise the symmetry which certain periods of relative minima display.

THEOREM 2.26

Let $w = \sqrt{\gamma/h}$. Then W_k can be represented as

$$W_k = (\theta_k(1,1) + h\, g_k W)/h\, N_\delta(A_k)$$

with θ_k, $hN_\delta(A_k) \in Z(\delta)$. Furthermore we can choose θ_k such that

$$|\theta_k| \leq (|\delta|/c + g_k)h|N_\delta(A_k)|/g_k$$

PROOF

$$W_k = B_k/A_k$$

$$= \left(\frac{\kappa_k+\lambda_k w}{\alpha_k+\beta_k w}, \frac{\kappa_k-\lambda_k w}{\alpha_k-\beta_k w}\right)$$

$$= ((\kappa_k+\lambda_k w)(\alpha_k-\beta_k w), (\kappa_k-\lambda_k w)(\alpha_k+\beta_k w))/N_\delta(A_k)$$

$$= (\kappa_k\alpha_k-\beta_k\lambda_k\gamma/h + (\alpha_k\lambda_k-\beta_k\kappa_k)w,$$

$$\kappa_k\alpha_k-\beta_k\lambda_k\gamma/h - (\alpha_k\lambda_k-\beta_k\kappa_k)w)/N_\delta(A_k)$$

$$= (((h\kappa_k\alpha_k-\beta_k\lambda_k\gamma)/h)(1,1) + g_k W)/N_\delta(A_k)$$

If we now set $\theta_k = h\kappa_k\alpha_k - \beta_k\lambda_k\gamma$ then the first result of the theorem is clear.

To show that we can choose θ_k to satisfy the stated bound we recall the discussion following theorem 2.20. There we saw that we can always replace W_k with $W_k + (\tau/g_k)(1,1)$, $\tau \in I_k'$ (equivalently replace θ_k with $\theta_k + hN_\delta(A_k)\tau/g_k$) thereby obtaining W_k satisfying

$$|W_k + W_k^*| < 2(g_k + |\delta|/c)/g_k$$

Since

$$|W_k + W_k^*| = 2|\theta_k/hN_\delta(A_k)|$$

the bound now follows. //

THEOREM 2.27

M(w) has periodic relative minima

if and only if

$w = \sqrt{\gamma/h}$, $\gamma \in Z(\delta)$, $h \in Z^+$, γ/h not a perfect square in $Q(\delta)$.

PROOF

The forward implication is just theorem 2.23. To prove the reverse implication we show that if $w = \sqrt{\gamma/h}$ then there exist $m,n \in Z$, $m \neq n$ such that $M_m(w) = M_n(w)$. The result will then follow from theorem 2.22b.

Therefore assume $w = \sqrt{\gamma/h}$. By theorem 2.16 we know that $M_k(w)$ is representable as

$$M_k(w) = \{(\alpha(1,1)+\beta W_k)/g_k : (\alpha,\beta)\ W_k \text{ allowable}\}$$

From the corollary to theorem 2.16 we see that g_k can only assume finitely many values. Consequently σ_k,ψ_k can only assume finitely many values. Furthermore if θ_k is required to satisfy the bound in theorem 2.26 then W_k can only assume finitely many values since

$\theta_k, hN_\delta(A_k)$, g_k are elements of bounded magnitudes in the discrete set $Z(\delta)$. Consequently the number of distinct $M_k(w)$ must be finite. However since $w \notin Q(\delta)$ it follows that $M(w)$ has an infinite chain of relative minima. Therefore we must find $M_m(w) = M_n(w)$ for some $m,n \in Z$ with $m \neq n$. //

We now consider some of the properties of and the relationships that exist between the A_k, $M_k(\sqrt{\gamma/h})$ corresponding to a period of relative minima of $M(\sqrt{\gamma/h})$. In particular we look at units, and the symmetry exhibited by certain minimal periods of relative minima. (These points have already been partially illustrated in examples 2.1, 2.3). The results obtained will be used in developing the modified version of algorithm 2.1 which we will develop in this section.

We begin with the subject of units. In examples 2.1, 2.3 it was perfectly clear what was meant by the term unit since $M(w)$ was closed under multiplication and hence formed a ring. More generally we find that $M(\sqrt{\gamma/h})$ is closed under multiplication if and only if $h = 1$. Consequently we need to define exactly what we mean by the term unit in order to clarify the situation for $h > 1$.

DEFINITION 2.13

Let $w = \sqrt{\gamma/h}$ and $U \in M(w)$. We say that U is a *unit* of $M(w)$ if and only if $M(w) = M(w)/U$. Furthermore we say that U *corresponds* to a unit of $Z(\delta)(w)$ if and only if the components of U (which are conjugates in $Q(\delta)(w)$) are units of $Z(\delta)(w)$. //

Note that requiring $M(w) = M(w)/U$ is equivalent to requiring $M(w)$ closed under multiplication by U, and $U^{-1} \in M(w)$. It is easily checked that if U,V are units of $M(w)$ then U^{-1}, U^*, UV are also units of $M(w)$.

EXAMPLE 2.6

(a) Let $\xi \in Z(\delta)$ be a root of unity. Then $U = (\xi,\xi)$ is a unit of $M(w)$ which corresponds to the unit $\xi \in Z(\delta)(w)$.

(b) Let $w = \sqrt{-1/2}$, $\delta = \sqrt{-2}$. Then $U = 2W = (\sqrt{-1},-\sqrt{-1}) \in M(w)$. Clearly U corresponds to the unit $\sqrt{-1} \in Z(\delta)(w)$. However since $UW = (-1/2,-1/2) \notin M(w)$ it is clear that $M(w)$ is not closed under multiplication by U. Consequently U is not a unit of $M(w)$.

(c) Recalling example 2.1 we see that

$$A_6 = (245-20\delta)(1,1) + (88-34\delta)W$$

is a unit of $M(\sqrt{1+\delta})$. Furthermore A_6 corresponds to the unit $(245-20\delta) + (88-34\delta)\sqrt{1+\delta} \in Z(\delta)(\sqrt{1+\delta})$. //

We now give the most important result concerning units of $M(w)$.

THEOREM 2.28

Let $w = \sqrt{Y/h}$ and $A,B \in M(w)$ satisfy $|AA^*|, |BB^*| \neq 0$. Then

$$M(w)/A = M(w)/B$$

if and only if

$$A/B, B/A \text{ are units of } M(w)$$

Furthermore if U is a unit of $M(w)$ then U corresponds to a unit of $Z(\delta)(w)$. In particular if $A_0,A_1,\ldots,A_q$ is a period of relative minima of $M(w)$ then A_q/A_0 is a unit of $M(w)$ which corresponds to a unit of $Z(\delta)(w)$.

PROOF

We have

$$M(w)/A = M(w)/B$$

if and only if

$$M(w) = M(w)/(B/A) = M(w)/(A/B)$$

Since $A/B = (1,1)/(B/A) \in M(w)/(B/A) = M(w)$ the first result of the theorem is clear.

Suppose U is a unit of M(w). Since $U \in M(w)$ we have $U = \alpha(1,1) + \beta W$ with $\alpha,\beta \in Z(\delta)$. Now $N_\delta(U)(1,1) = UU^* \in M(w)$ since $U^* \in M(w)$ and M(w) is closed under multiplication by U. Consequently $N_\delta(U) \in Z(\delta)$. However U^{-1} is also a unit of M(w) and so we also have $N_\delta(U^{-1}) = 1/N_\delta(U) \in Z(\delta)$. Thus $N_\delta(U)$ is a $Z(\delta)$ root of unity and it therefore follows that the components of U, that is $\alpha \pm \beta w$, are units of $Z(\delta)(w)$.

The final result of the theorem follows from the first two since $M_0(w) = M_q(w)$. //

In view of this theorem we see that we can obtain a unit of $Z(\delta)(w)$ by calculating a period of relative minima of M(w). Unfortunately in contrast with the standard quadratic/simple continued fraction case, it is not always possible to choose w such that the unit obtained is guaranteed to be a fundamental unit of $Z(\delta)(w)$. In chapter three we will show how this problem can be overcome.

Of course A_q/A_0 is not the only unit of M(w). In fact if $\{A_k : k \in Z\}$ is a complete chain of relative minima of M(w) satisfying $M_k(w) = M_{k+q}(w)$ $\forall$ $k \in Z$ (see theorem 2.22c) then A_{k+mq}/A_k is also a unit of M(w) for any $k,m \in Z$.

Although the proof of theorem 2.22c indicates one way of defining a complete chain from a period it is simpler to use a generalization of the method used in example 2.1 which is given in our next theorem.

THEOREM 2.29

Let $w = \sqrt{Y/h}$.

(a) If U is a unit of M(w) and A_k is a relative minimum of M(w) then UA_k is a relative minimum of M(w).

(b) Suppose $A_0, A_1, \ldots, A_q$ is a period of relative minima of M(w). A complete chain of relative minima of M(w) is given by

$$A_{k+mq} = A_k(A_q/A_0)^m,\ m \in Z,\ k = 0,1,\ldots,q-1$$

PROOF

A generalization of the argument used in example 2.1 proves the result. //

We leave further discussion of the subject of units until chapter three where we consider in detail the problem of calculating fundamental units of $Z(\delta)(\sqrt{Y})$.

We now move on to consider the symmetry exhibited by certain periods of relative minima. As a consequence of the results developed we will be able to reduce the amount of work required to calculate a period of relative minima. To simplify matters we shall restrict our attention to minimal periods, that is periods of minimal length, which in practice are the only periods of interest from an efficiency point of view. (However it will be fairly obvious that the following results also apply in a modified form to non-minimal periods).

We begin with a definition.

DEFINITION 2.14

Let $M \subseteq \mathbb{C} X \mathbb{C}$. Then we define

$$M^* = \{R^* : R \in M\}$$

For simplicity we shall write $M^*(w)$, $M_k^*(w)$ rather than $(M(w))^*$, $(M_k(w))^*$ //

Clearly $M(w) = M^*(w)$ and it follows that

$$M_k^*(w) = M^*(w)/A_k^* = M(w)/A_k^*$$

The modules $M_k^*(w)$ are used in our next theorem to describe the symmetry which occurs for certain periods of relative minima. Note that the symmetries which occur are very similar to those which occur in the simple continued fraction expansion of $\sqrt{d_1}$, d_1 a square-free positive integer.

THEOREM 2.30

Let $w = \sqrt{\gamma/h}$. We can choose a minimal period of relative minima $A_0, A_1, .., A_q$ with $A_1, A_2, .., A_{[q/2]}$ arbitrary such that the following properties hold.

(a) If q is odd then

$$M_{q-k}(w) = M_k^*(w), \quad N_\delta(A_{q-k}) = \xi N_\delta(A_k) \tag{14}$$

for $k = 0,1,\ldots,(q-1)/2$ where $\xi = N_\delta(A_q/A_0)$ is a $Z(\delta)$ root of unity.

(b) If q is even then (14) holds for $k = 0,1,\ldots,q/2-1$ and we can find $A_{q/2}^{(1)}$ such that

$$M_{q/2}^{(1)}(w) = M_{q/2}^*(w), \quad N_\delta(A_{q/2}^{(1)}) = \xi N_\delta(A_{q/2})$$

PROOF

Let q be the length of a minimal period of M(w). Thus there exists A_q such that $M_0(w) = M_q(w)$. Choose $A_1, A_2, \ldots, A_{[q/2]}$ arbitrarily. We will show that the required period can be constructed from $A_0, A_1, \ldots, A_{[q/2]}$ and A_q.

Since A_q/A_0 is a unit of M(w) we have from theorem 2.29a that $A_k^* A_q/A_0$, $k = 1,2,\ldots,[q/2]$ are relative minima of M(w). By considering magnitudes it is a simple matter to check that $A_k^* A_q/A_0 \in E_{q-k}$.

(a) Suppose q is odd. We have $[q/2] = (q-1)/2$ and so setting

$$A_{q-k} = A_k^* A_q/A_0, \quad k = 1,2,\ldots,(q-1)/2$$

completes the construction of a minimal period. Since A_q/A_0 is a unit

of $M(w)$ we have $M_{q-k}(w) = M_k^*(w)$ and the norm relationship is obvious once we recall that $N_\delta(A) = N_\delta(A^*)$.

(b) Suppose q is even. We have $[q/2] = q/2$ and so setting

$$A_{q-k} = A_k^* A_q/A_0, \quad k = 1,2,\ldots,q/2-1$$

$$A_{q/2}^{(1)} = A_{q/2}^* A_q/A_0$$

completes the construction of the required minimal period plus $A_{q/2}^{(1)}$. //

Theorem 2.30 guarantees that in any $M(\sqrt{\gamma/h})$ there exist periods which we will describe as symmetric minimal periods. Our interest in such periods lies in the fact that we can take advantage of their existence when calculating a period of relative minima. Note that the proof of theorem 2.30 shows that a symmetric minimal period can be generated from A_0, arbitrary $A_1, A_2, .., A_{[q/2]}$, plus an appropriate A_q. Since $A_{(q+1)/2} = A_{(q-1)/2}^* A_q/A_0$ (q odd) and $A_{q/2}^{(1)} = A_{q/2}^* A_q/A_0$ (q even) we can take the alternative view that a symmetric minimal period can be generated from A_0, arbitrary $A_1, A_2, \ldots, A_{[q/2]}$, plus an appropriately chosen $A_{(q+1)/2}$ (q odd) or $A_{q/2}^{(1)}$ (q even). Thus if as we calculate a chain $A_0, A_1, \ldots, A_k$ we can easily locate the appropriate $A_{(q+1)/2}$ (q odd) or $A_{q/2}^{(1)}$ (q even) when we reach $k = [q/2]$ then we can expect to approximately halve the amount of work involved in calculating a minimal period.

EXAMPLE 2.7

(a) Recall example 2.3. Instead of calculating a period directly we can proceed as follows.

Calculate A_0, A_1, A_2, A_3 at which point we note that $M_3(w) = M_3^*(w)$. (This can be verified algebraically as noted in example 2.3. However we shall develop a simpler testing procedure in theorem 2.32). Thus A_3/A_3^* is a unit of $M(w)$ and so the chain A_0, A_1, A_2, A_3 can be extended

into a period of length 6 by setting

$$A_4 = A_2^*(A_3/A_3^*),\ A_5 = A_1^*(A_3/A_3^*),\ A_6 = A_3/A_3^*$$

If we take $A_3^{(1)} = A_3$ then the symmetries in theorem 2.30b clearly hold. ($\xi = N_\delta(A_6) = 1$). The fact that this period is minimal can be deduced from the fact that $|N_\delta(A_k/A_0)| \neq 1$ for $0 < k < 6$. (See example 2.1 for the $N_\delta(A_k)$).

Of particular interest is the fact that we can take $A_3^{(1)} = A_3$. More generally we have found that in practice we have always been able to take $A_{q/2}^{(1)} = A_{q/2}$ when q is even. However we have been unable to prove this will always be the case.

(b) Let $\delta = \sqrt{-35}$, $\gamma = -3+5\delta$, $h = 1$, $w = \sqrt{\gamma/h}$. Algorithm 2.1 produces the following results.

TABLE 2.11 Relative minima of $M(\sqrt{-3+5\delta})$

k	α_k	β_k	$N_\delta(A_k)$
0	1	0	1
1	$(15+3\delta)/2$	2	$(-21+5\delta)/2$
2	$(27-3\delta)/2$	$(1-\delta)/2$	$(-19+\delta)/2$
3	$-43-4\delta$	$(-17+\delta)/2$	$-8+\delta$
4	$(-255-7\delta)/2$	$(-37+5\delta)/2$	$(21-5\delta)/2$
5	$62-25\delta$	$(-25-9\delta)/2$	-1

Clearly A_5 is a unit of M(w) and we have $M_0(w) = M(w) = M_5(w)$. Thus $A_0, A_1, \ldots, A_5$ is a period which is easily seen to be minimal. However this minimal period does not satisfy all the conditions in theorem 2.30a. To be more precise we do not have $M_3(w) = M_2^*(w)$, nor do we have $N_\delta(A_3) = \xi N_\delta(A_2)$ where ξ is a $Z(\delta)$ root of unity. (Note however that $|N_\delta(A_3)| = |N_\delta(A_2)|$). Consequently no period "midpoint" was recognised during the calculation of the chain $A_0, A_1, \ldots, A_5$, that is there was no occurrence of $M_{k+1}(w) = M_k^*(w)$.

Of course the problem is that algorithm 2.1 only produces one $A_k \in E_k$. Consequently

$$A_3^{(1)} = A_2^* A_5 = ((-61+13\delta)/2)(1,1) + ((3+3\delta)/2)W$$

was not produced for consideration. It is a simple matter to check that the minimal period $A_0, A_1, A_2, A_3^{(1)}, A_4, A_5$ exhibits the symmetry described in theorem 2.30a. $(\xi = N_\delta(A_5) = -1)$.

As a final point we note that $A_3^{(1)}/A_3 = 3/((1+\delta)/2)$, $|(1+\delta)/2| = 3$, and 9 factors nonuniquely in $Z(\delta)$ as

$$9 = 3.3 = ((1+\delta)/2)((1-\delta)/2) \qquad //$$

Example 2.7 clearly indicates the need for a more systematic approach to the problem of locating the midpoint of a symmetric minimal period. The next theorem suggests the basic outline for such a procedure.

THEOREM 2.31

Let $w = \sqrt{Y/h}$ and suppose $A_0, A_1, \ldots, A_k$ is a chain of relative minima of $M(w)$ satisfying

$$M_{j+1}^{(m)}(w) \neq M_j^*(w), \quad M_{j+1}^{(m)}(w) \neq M_{j+1}^*(w), \quad j = 0,1,2,..,k-1$$

where $A_{j+1}^{(m)}$ runs through all elements of E_{j+1}.

(a) If

$$\exists\ A_{k+1} \in E_{k+1} \text{ such that } M_{k+1}(w) = M_k^*(w) \tag{15}$$

then $A_0, A_1, \ldots, A_k$ plus $A_q = (A_{k+1}/A_k^*)A_0$ generate a minimal period of odd length $q = 2k+1$ as described in the proof of theorem 2.30a.

(b) If (15) does not hold and

$$\exists\ A_{k+1}, A_{k+1}^{(1)} \in E_{k+1} \ (A_{k+1} \text{ arbitrary}) \text{ such that } M_{k+1}^{(1)}(w) = M_{k+1}^*(w) \tag{16}$$

then $A_0, A_1, \ldots, A_k, A_{k+1}$ plus $A_q = (A_{k+1}^{(1)}/A_{k+1}^*)A_0$ generate a minimal period of even length $q = 2k+2$ as described in the proof of theorem 2.30b.

PROOF

(a) Set $U = A_{k+1}/A_k^*$. By theorem 2.29a we have UA_j^*, $j = 0,1,\ldots,k$ are relative minima of $M(w)$. Now

$$|UA_0^*| > |UA_1^*| > \ldots > |UA_k^*| = |A_{k+1}|$$

It is therefore easily deduced that $UA_j^* \in E_{r-j}$ where $r = 2k+1$. In particular $UA_0^* = (A_{k+1}/A_k^*)A_0 \in E_r$ and so we can choose $A_r = (A_{k+1}/A_k^*)A_0$. Clearly $M_0(w) = M_r(w)$. To complete the proof we need only show that $r = q$ where q is the length of a minimal period. Clearly $q|r$ and so q is odd. Now theorem 2.30a shows that $M(w)$ has a symmetric minimal period

$$A_0, A_1^{(1)}, \ldots, A_{(q-1)/2}^{(1)}, A_{(q+1)/2}^{(1)}, \ldots, A_q^{(1)}$$

with $A_{(q-1)/2}^{(1)}$ arbitrary and $M_{(q+1)/2}^{(1)}(w) = (M_{(q-1)/2}^{(1)}(w))^*$. Since $(q-1)/2 \leq (r-1)/2 = k$ we can choose $A_{(q-1)/2}^{(1)} = A_{(q-1)/2}$. Thus there exists $A_{(q+1)/2}^{(1)}$ such that $M_{(q+1)/2}^{(1)}(w) = M_{(q-1)/2}^*(w)$. Now if $(q-1)/2 < k$ then the conditions in the theorem are contradicted. Therefore $(q-1)/2 = k$ and $q = 2k+1 = r$.

(b) The proof of this part is similar to part (a). The assumption that (15) does not hold is included to avoid an incorrect conclusion when A_0, A_1 is a minimal period. (Part (b) would otherwise incorrectly conclude that A_0, A_1, A_2 is a minimal period). The only other difference is that in the second half of the proof we must consider the case q even as well as q odd. //

In view of this theorem (plus the obvious correspondence between $A_{k+1}^{(m)} \in M(w)$ and $R_k^{(m)} \in M_k(w)$) it is fairly easily seen that the following modified version of algorithm 2.1 will calculate a chain $A_0, A_1, \ldots \in M(w)$ and stop when the midpoint of a minimal period is reached.

(i) Choose A_0 and set $k = 0$

(ii) Calculate all $R_k \in M_k(w)$ up to $Z(\delta)$ root of unity factors. Denote these R_k by $R_k^{(1)}, R_k^{(2)}, \ldots, R_k^{(n)}$
(It suffices to consider the distinct R_k up to $Z(\delta)$ root of unity factors since if $R_k^{(1)} = \xi R_k$ (equivalently $A_{k+1}^{(1)} = \xi A_{k+1}$) with ξ a $Z(\delta)$ root of unity then $M_{k+1}^{(1)}(w) = M_{k+1}(w)$).

(iii) Test to see if any $R_k^{(j)}$, $j = 1,2,\ldots,n$ gives

$$M_{k+1}^{(j)}(w) = M_k^*(w)$$

If any test is positive then stop - the midpoint of a minimal period of odd length has been found. (This step effectively tests for an occurrence of (15).

(iv) Test to see if any $R_k^{(j)}$, $j = 1,2,\ldots,n$ gives

$$M_{k+1}^{(j)}(w) = (M_{k+1}^{(1)}(w))^*$$

If any test is positive then stop - the midpoint of a minimal period of even length has been found. (This step effectively tests for an occurrence of (16). The fixed right hand side of the module equation reflects the fact that A_{k+1} in (16) can be selected arbitrarily).

(v) Set $A_{k+1} = R_k^{(1)} A_k$, increment k, go to (ii).

Of course the above procedure is only a brief outline of what will eventually become algorithm 2.2, that is the modified version of algorithm 2.1 for calculating periodic relative minima. In the following paragraphs we discuss the necessary finer details of algorithm 2.2 which have not already been covered in the discussion of algorithm 2.1.

Note that although algorithm 2.1A only produces one $R_k \in M_k(w)$ it clearly considers and then at step 11 discards the remaining $R_k^{(j)}$ required in step (ii) above. Furthermore we generally find $n = 1$ in step (ii) and it is unusual to find $n > 2$. Thus step (ii) involves very little work that is extra to that already required by algorithm 2.1A.

The main point left to consider is the development of a simple method for the testing required in (iii) and (iv), that is a method for testing whether or not $M_m(w) = M_j^*(w)$.

THEOREM 2.32

Let $w = \sqrt{\gamma/h}$ and suppose A_j, A_m are relative minima of $M(w)$ with $M_j(w)$, $M_m(w)$ represented as in theorem 2.16 and W_j, W_m represented as in theorem 2.26. Then

$$M_m(w) = M_j^*(w)$$

if and only if

(a) $N_\delta(A_m) = \xi N_\delta(A_j)$ where ξ is a $Z(\delta)$ root of unity

(b) $g_m = g_j$

(c) $\sigma_m = \sigma_j$

(d) $g_m(\theta_m + \theta_j)/h\, N_\delta(A_m) = \lambda \in I_m'$

(e) $\psi_m + \sigma_m \lambda / g_m + \xi^{-1}\psi_j \in I_m'$

PROOF

Suppose $M_m(w) = M_j^*(w)$. Since A_m/A_j^* is a unit of $M(w)$ we have $\xi = N_\delta(A_m/A_j^*) = N_\delta(A_m)/N_\delta(A_j)$ is a $Z(\delta)$ root of unity. Now a representation of $M_j^*(w)$ can be obtained from a representation of $M_j(w)$ by simply replacing W_j with W_j^*. (The representation is not standard in the sense of theorem 2.16 since $W_j^* = A_j^*/B_j^*$, $A_j^* = \alpha_j(1,1) + (-\beta_j)W$, $B_j^* = \kappa_j(1,1) + (-\lambda_j)W$ and so $\alpha_j(-\lambda_j) - (-\beta_j)\kappa_j = -g_j$). Since $(\sigma_m'/g_m)(1,1) \in M_m(w) = M_j^*(w)$ we have $(\sigma_m'/g_m)(1,1) = (\alpha(1,1) + \beta W_j^*)/g_j$ with

(α,β) W_j allowable. Let $\alpha = (a+b\delta)/c$. Now $(1,1)$, W_j are linearly independent and so it follows that $\beta = 0$, $\sigma_m'/g_m = \alpha/g_j$. The imaginary components of this equality give $-1/cg_m = b/cg_j$. Consequently $g_m | g_j$. A similar argument gives $g_j | g_m$ and so $g_m = g_j$. It then follows easily that $\sigma_m = \sigma_j$. To prove (d) we note that $W_m, W_j^* \in M_m(w) = M_j^*(w)$ and so $W_m + \xi^{-1}W_j^* \in M_m(w)$. ($\xi$ is as defined by (a)). Now

$$W_m + \xi^{-1}W_j^* = (\theta_m(1,1) + hg_mW)/hN_\delta(A_m) + (\theta_j(1,1) - hg_jW)/\xi hN_\delta(A_j)$$

$$= ((\theta_m+\theta_j)/hN_\delta(A_m))(1,1)$$

and so $(\theta_m+\theta_j)/hN_\delta(A_m) = \lambda/g_m$ with $\lambda \in I_m'$. (The representation of $M_m(w)$ which can be derived from $M_j(w)$ via the relationship $M_m(w) = M_j^*(w)$ is standard if and only if $\xi = -1$. Consequently the relationship between W_m, W_j^* is not quite what we might initially expect given the result of theorem 2.20 which only considered standard representations). Finally we note that

$$(\psi_m(1,1) + \sigma_m W_m)/g_m = (\psi_m(1,1) + \sigma_m(-\xi^{-1}W_j^* + \lambda(1,1)/g_m))/g_m$$

$$= (\psi_m + \sigma_m\lambda/g_m + \xi^{-1}\psi_j)/g_m - \xi^{-1}(\psi_j(1,1) + \sigma_jW_j^*)/g_j$$

Now $(\psi_m(1,1) + \sigma_mW_m)/g_m$, $(\psi_j(1,1) + \sigma_jW_j^*)/g_j \in M_m(w) = M_j^*(w)$ and it therefore follows that $((\psi_m + \sigma_m\lambda/g_m + \xi^{-1}\psi_j)/g_m)(1,1) \in M_m(w)$ which gives (e).

The reverse implication is straight forward. Briefly, if $R = (\alpha(1,1) + \beta W_m)/g_m \in M_m(w)$ then using the notation of theorem 2.16 plus (a) - (e) we have (note that (a),(d) give $W_m = -\xi^{-1}W_j^* + (\lambda/g_m)(1,1)$)

$$R = ((m_1g_m + m_2\sigma_m' + m_4\psi_m)(1,1) + (m_3g_m + m_4\sigma_m)W_m)/g_m$$

$$= (m_1g_m + m_2\sigma_m' + m_3\lambda + m_4(\psi_m + \sigma_m\lambda/g_m + \xi^{-1}\psi_j))(1,1)/g_j$$

$$-m_3\xi^{-1}W_j^* - \xi^{-1}m_4(\psi_j(1,1) + \sigma_jW_j^*)/g_j$$

$$\in M_j^*(w)$$

Similarly $R = (\alpha(1,1) + \beta W_j^*)/g_j \in M_j^*(w)$ implies $R \in M_m(w)$. //

Note that if $A_j = A_m$ then conditions (a), (b), (c) are automatically satisfied while condition (d) reduces to $2g_m\theta_m/hN_\delta(A_m) = \lambda \in I_m'$ and condition (e) reduces to $2\psi_m + \sigma_m\lambda/g_m \in I_m'$.

We now consider how best to use conditions (a) - (e) to carry out the testing required in steps (iii), (iv) of the procedure following theorem 2.31. Note that the calculation of an $M_{k+1}^{(j)}(w)$ representation is a non-trivial matter. Consequently we use condition (a) as follows to eliminate as many cases as possible before any $M_{k+1}^{(j)}(w)$ representation is calculated. Begin by relabelling the $R_k^{(j)}$ if possible so that $N_\delta(R_k^{(1)})$ is a root of unity and then relabel the $R_k^{(j)}$, $j = 2,\ldots,n$ so that $N_\delta(R_k^{(j)}/R_k^{(1)})$ is a root of unity for $j = 1,\ldots,m$ but not for $j = m+1,\ldots,n$. If (15) is to occur then we must have $N_\delta(R_k^{(j)}) = N_\delta(A_{k+1}^{(j)}/A_k)$ a $Z(\delta)$ root of unity. Therefore depending on $N_\delta(R_k^{(1)})$ we will have either 0 or m possibilities which require further consideration. (That is conditions (b) - (e) must be considered which requires the calculation of an $M_{k+1}^{(j)}(w)$ representation). If (16) is to occur then we must have $N_\delta(A_{k+1}^{(j)}/A_{k+1}^{(1)}) = N_\delta(R_k^{(j)}/R_k^{(1)})$ a $Z(\delta)$ root of unity and so there are m possibilities which require further consideration. Note that the $R_k^{(j)}$, $j = m+1,\ldots,n$ are not considered again after the relabelling process and so they can be discarded.

In spite of the apparent complexity of the testing outlined in the previous paragraph the situation is quite simple in practice. In general we find that after relabelling we have $m = 1$ and $N_\delta(R_k^{(1)})$ is not a root of unity. (In fact we often find $n = 1$ so we do not even have to consider relabelling). Thus in general the only case which requires any consideration is the possibility $M_{k+1}^{(1)}(w) = (M_{k+1}^{(1)}(w))^*$. Consequently the testing for a midpoint normally only requires a trivial amount of

work which is extra to that already required by algorithm 2.1. The only situation where the required testing would involve a significant amount of work is where $m > 1$ since we would then have to calculate at least two $M_{k+1}^{(j)}(w)$ representations. Fortunately such occurrences appear to be rare. (Indeed we have not found any).

We now collect together the ideas and results presented in this section to form algorithm 2.2. We begin by making the necessary modifications to algorithms 2.1A, 2.1B. The resulting algorithms 2.2A, 2.2B form the major parts of algorithm 2.2.

ALGORITHM 2.2A

Given a standard representation of $M_k(w)$ we calculate all $R_k \in M_k(w)$ up to $Z(\delta)$ root of unity factors as follows.

1 - 10 are the same as in algorithm 2.1A

11 If $r > mR$, or $r = mR$ and $r^* > mR^*$ go to 13

12.1 If $rr^* < mR\, mR^*$ then set $n = 1$, $mR = r$, $mR^* = r^*$, $b(\beta) = g_k(mR+1)/|w_{k,1} - w_{k,2}|$, go to 12.3

12.2 Reset $n = n+1$

12.3 Set $\eta_{k,1}^{(n)} = m_1 g_k + m_2 \sigma_k' + m_4 \psi_k$, $\eta_{k,2}^{(n)} = \beta$, $R_k^{(n)} = (\eta_{k,1}^{(n)}(1,1) + \eta_{k,2}^{(n)} W_k)/g_k$

13-23 are the same as in algorithm 2.1A

24 Stop //

ALGORITHM 2.2B

Given a standard representation of $M_k(w)$ and $R_k^{(j)} = (\eta_{k,1}^{(j)}(1,1) + \eta_{k,2}^{(j)} W_k)/g_k$ we calculate a standard representation of $M_{k+1}^{(j)}(w)$ as follows.

1 is the same as in algorithm 2.1B

2 Set $\eta_1 = -\eta_{k,2}^{(j)}$, $\eta_2 = -\sigma_k' \eta_{k,2}^{(j)}/g_k$, $\eta_3 = \eta_{k,1}^{(j)}$, $\eta_4 = (\sigma_k \eta_{k,1}^{(j)} - \psi_k \eta_{k,2}^{(j)})/g_k$

3-30 are the same as in algorithm 2.1B

31 Set $g_{k+1}^{(j)} = g_{k+1}$, $\sigma_{k+1}^{(j)} = \sigma_{k+1}$, $\psi_{k+1}^{(j)} = \psi_{k+1}$, $W_{k+1}^{(j)} = W_{k+1}$

32 Set $hN_\delta(A_{k+1}^{(j)}) = N_\delta(R_k^{(j)}) h N_\delta(A_k)$

33 Set $\theta_{k+1}^{(j)} = h N_\delta(A_{k+1}^{(j)}) (w_{k+1,1}+w_{k+1,2})/2$

34 Stop //

Thus the only changes are the addition of superscripts to distinguish the various representations and the calculation of $\theta_{k+1}^{(j)}$, $h N_\delta(A_{k+1}^{(j)})$ which are required to perform the tests listed in theorem 2.32. Note that it is more convenient to work with the $Z(\delta)$ integer $h N_\delta(A_k)$ rather than $N_\delta(A_k)$ which is not necessarily integral when $h > 1$.

ALGORITHM 2.2

Let $w = \sqrt{\gamma/h}$. A symmetric minimal period of relative minima of $M(w)$ can be calculated as follows.

1 If $|w| \geq 1$ set $A_0 = (1,1)$, $h N_\delta(A_0) = h$, $W_0 = W$, go to 3

2 Set $A_0 = W$, $h N_\delta(A_0) = -\gamma$, $W_0 = -W^{-1}$

3 Set $g_0 = 1$, $\psi_0 = 0$, $\theta_0 = 0$, $\sigma_0 = \omega$

4 Set $k = 0$

5 Use algorithm 2.2A to calculate $R_k^{(j)}$, $j = 1,2,\ldots,n$ (that is all $R_k \in M_k(w)$ up to $Z(\delta)$ root of unity factors)

6 If $|N_\delta(R_k^{(1)})| \neq 1$ go to 8

7 If $N_\delta(R_k^{(j)}) = \xi$ occurs then relabel so that $N_\delta(R_k^{(1)}) = \xi$

8 Discard all $R_k^{(j)}$, $j > 1$ for which $N_\delta(R_k^{(j)}/R_k^{(1)}) \neq \xi_1$, relabel the remaining $R_k^{(j)}$ and reset n accordingly

9 For $j = 1,2,\ldots,n$ use algorithm 2.2B to calculate the representation of $M_{k+1}^{(j)}(w)$ corresponding to $R_k^{(j)}$

10 If $N_\delta(R_k^{(1)}) \neq \xi$ go to 13

11 For $j = 1,2,\ldots,n$ perform step 12

12 If $M_{k+1}^{(j)}(w) = M_k^*(w)$ then set $q = 2k+1$, $A_{k+1}^{(j)} = R_k^{(j)} A_k$, go to 17

13 Set $A_{k+1}^{(1)} = R_k^{(1)} A_k$

14 For $j = 1,2,\ldots,n$ perform step 15

15 If $M_{k+1}^{(j)}(w) = (M_{k+1}^{(1)}(w))^*$ then set $q = 2k+2$, $A_{k+1}^{(j)} = R_k^{(j)} A_k$, go to 17

16 Set $A_{k+1} = A_{k+1}^{(1)}$, $M_{k+1}(w) = M_{k+1}^{(1)}(w)$, increment k by 1, go to 5

17 If q is odd then $A_0, A_1, \ldots, A_k, A_{k+1}^{(j)}$ define a symmetric minimal period as described in theorem 2.31a

else $A_0, A_1, \ldots, A_k, A_{k+1}^{(1)}, A_{k+1}^{(j)}$ define a symmetric minimal period as described in theorem 2.31b

18 Stop

Explanatory Notes

(1) ξ, ξ_1 in steps 7, 8, 10 denote arbitrary $Z(\delta)$ roots of unity.

(2) Steps 6, 7, 8 perform initial testing for a period midpoint using condition (a) of theorem 2.32 as described in the paragraphs following that theorem. Note $|N_\delta(R_k^{(j)})| = |N_\delta(R_k^{(1)})|$ and so we must have $|N_\delta(R_k^{(1)})| = 1$ if $N_\delta(R_k^{(j)}) = \xi$ is to occur.

(3) The tests required in step 12 are given by (b), (c), (d), (e) of theorem 2.32 and the tests required in step 15 are given by (d), (e) of theorem 2.32. //

Before illustrating algorithm 2.2 with several examples we look at some important practical aspects of implementing the algorithm. In the paragraph following algorithm 2.1 we noted that the precision of calculation required by the algorithm depends on the maximum magnitude of the relative minima which we wish to calculate. Consequently we noted that multiprecision arithmetic would be required for the calculation of relative minima of large magnitude. There are two areas

in which the precision of the calculation is important. Firstly the calculation of W_{k+1} and secondly the calculation of $A_{k+1},\alpha_{k+1},\beta_{k+1}$. In the following paragraphs we show that for algorithm 2.2 we can avoid the need for multiprecision arithmetic in the calculation of W_{k+1}. (To be more precise we shall show that the precision required depends on $|\gamma|$, h, $|\delta|$ rather than $|A_{k+1}|$). Furthermore we can reduce the amount of computation required to calculate $A_{k+1},\alpha_{k+1},\beta_{k+1}$ when $|A_{k+1}|$ is large.

The need for multiprecision arithmetic in the calculation of W_{k+1} can be eliminated by making the following simple modifications to algorithm 2.2. Firstly we assume that the calculations in steps 31 and 32 of algorithm 2.2B involve rounding the final result to the nearest $Z(\delta)$ integer. Secondly we add the following calculation to step 16 of algorithm 2.2

$$\text{Reset } W_{k+1} = (\theta_{k+1}(1,1) + hg_{k+1}W)/hN_{\delta}(A_{k+1})$$

Now theorems 2.25, 2.26 show that $\theta_{k+1}, hN_{\delta}(A_{k+1})$ are "small" $Z(\delta)$ integers (bounded independent of k but depending on γ,h,δ) and we have $|\alpha-\beta| \geq 1$ when $\alpha,\beta \in Z(\delta), \alpha \neq \beta$. It is therefore not difficult to see that we only require a fairly low level of precision in our calculations (depending on $|\gamma|,h,|\delta|$) in order for the above modifications to ensure that the calculated values of θ_{k+1}, $hN_{\delta}(A_{k+1})$, W_{k+1} are always correct to within rounding error regardless of how large k grows. (Standard precision (that is approximately 10S) is more than sufficient for the examples which we will consider). In effect the algorithm becomes self correcting and we are able to calculate $R_0, R_1, \ldots, R_k$ for arbitrarily large k.

Note that in making these modifications we are simply using the basic idea behind the simple continued fraction algorithm for quadratic surds. (See algorithm 1.2. For a generalization see Hendy and Jeans [1981]).

This basic idea is that we represent algebraic numbers (such as W_{k+1}) in terms of integral coefficients ($\theta_{k+1}, g_{k+1}, h N_\delta(A_{k+1})$ for W_{k+1}) which can be calculated exactly at each stage of the algorithm. Note that we could have followed this basic idea more closely by rewriting the calculation of W_{k+1} in terms of calculating $\theta_{k+1}, h N_\delta(A_{k+1})$ from $\theta_k, hN_\delta(A_k)$, etc using only $Z(\delta)$ arithmetic. However this proves to be a less efficient alternative to the modifications noted in the previous paragraph and so is not pursued.

Thus the only section of algorithm 2.2 which may still require multiprecision arithmetic is the calculation of $A_{k+1}, \alpha_{k+1}, \beta_{k+1}$. (That is steps 12,13,15,17 of algorithm 2.2). In some cases we only require a standard precision approximation to A_{k+1} (for example, class number calculations) and in such cases the algorithm has no need at all for multiprecision arithmetic. However if we wish to calculate the coefficients $\alpha_{k+1}, \beta_{k+1}$ or a high precision approximation to A_{k+1} then we cannot avoid the use of multiprecision arithmetic when $|A_{k+1}|$ is large.

Although the method for calculating $\alpha_{k+1}, \beta_{k+1}$ (from A_{k+1}) given in step 6 of algorithm 2.1 is quite satisfactory for small relative minima we can use an alternative method for large relative minima which uses only $Z(\delta)$ arithmetic. We derive this method as follows. The equation $A_{k+1} = R_k A_k$ is equivalent to

$$\alpha_{k+1}(1,1) + \beta_{k+1} W = ((\eta_{k,1}(1,1) + \eta_{k,2} W_k)/g_k)(\alpha_k(1,1) + \beta_k W)$$

Substituting $W_k = (\theta_k(1,1) + hg_k W)/h N_\delta(A_k)$ and expanding gives

$$\begin{aligned} \alpha_{k+1} &= (\alpha_k(hN_\delta(A_k)\eta_{k,1} + \eta_{k,2}\theta_k) + \beta_k \eta_{k,2} g_k \gamma)/hN_\delta(A_k) g_k \\ \beta_{k+1} &= (\beta_k(hN_\delta(A_k)\eta_{k,1} + \eta_{k,2}\theta_k) + \alpha_k \eta_{k,2} h g_k)/hN_\delta(A_k) g_k \end{aligned} \tag{17}$$

This provides a more efficient method for calculating $\alpha_{k+1}, \beta_{k+1}$ when

$|A_{k+1}|$ is large. The reasons for this are as follows. The method based on (17) only requires simple multiprecision integer arithmetic, that is addition of multiprecision integers and multiplication/division of multiprecision integers by standard precision integers. The amount of work involved is therefore a linear function of the precision. However the method used in algorithm 2.1 requires multiplication/division of multiprecision reals and the amount of work involved is therefore a quadratic function of the precision.

Once the midpoint of a period has been found we can calculate the coefficients of the remaining relative minima (step 17 of algorithm 2.2) by using the following forms. Let $A = \alpha(1,1) + \beta W$, $B = \kappa(1,1) + \lambda W$. Then

$$\begin{aligned} A/B &= AB^*/N_\delta(B) \\ &= ((\alpha\kappa - \beta\lambda\gamma/h)(1,1) + (-\alpha\lambda+\beta\kappa)W)/N_\delta(B) \end{aligned} \tag{18}$$

and

$$AB = (\alpha\kappa+\beta\lambda\gamma/h)(1,1) + (\alpha\lambda+\beta\kappa)W$$

Often we are only interested in calculating the coefficients of A_q. (See chapter three). In such cases we only require one calculation of the form (18) to give α_q, β_q once the mid point of a period has been reached.

The coefficient α_k also leads to a simple but generally satisfactory approximation to A_k when $|A_k|$ is large. Note that $|A_k - 2\alpha_k(1,1)| = |A_k^*|$. Thus $2\alpha_k$ is a good approximation to the first component of A_k while $N_\delta(A_k)/2\alpha_k$ is a good approximation to the second component of A_k when $|A_k|$ is large.

Thus although we cannot always avoid the use of multiprecision arithmetic when $|A_k|$ is large we can use the procedures outlined in the above paragraphs to reduce the amount of calculation involved thereby making the algorithm more efficient.

We now illustrate the use of algorithm 2.2 with several examples.

EXAMPLE 2.8

We rework example 2.1, 2.3 using algorithm 2.2. We have $\delta = \sqrt{-10}$, $\gamma = 1+\delta$, $h = 1$, $w = \sqrt{\gamma/h}$. The initialization steps give $A_0 = (1,1)$ and

$$g_0 = 1,\ \psi_0 = 0,\ \sigma_0 = 0,\ \theta_0 = hN_\delta(A_0) = 1,\ W_0 = W$$

For $k = 0$ we obtain the following results. Algorithm 2.2A gives

$$R_0^{(1)} = (4+\delta)(1,1) + 3W_0$$

Since $|N_\delta(R_0^{(1)})| \approx 4.3$ and $n = 1$ we can move straight to step 9 of algorithm 2.2 where we calculate a representation of $M_1^{(1)}(w)$ which is defined by

$$g_1^{(1)} = 1,\ \psi_1^{(1)} = 0,\ \sigma_1^{(1)} = \delta,\ \theta_1^{(1)} = 3-\delta,\ hN_\delta(A_1^{(1)}) = -3-\delta$$

(that is $W_1^{(1)} = ((3-\delta)(1,1) + W)/(-3-\delta)$)

Now $N_\delta(R_0^{(1)})$ is not a root of unity so $M_1^{(1)}(w) \neq M_0^*(w)$. We therefore set

$$A_1^{(1)} = R_0^{(1)}A_0 = (4+\delta)(1,1) + 3W$$

and then test to see if $M_1^{(1)}(w) = (M_1^{(1)}(w))^*$. Since

$$\lambda = 2g_1^{(1)}\theta_1^{(1)}/hN_\delta(A_1^{(1)}) \approx .11 + .63\delta \notin I_1'$$

it is clear that $M_1^{(1)}(w) \neq (M_1^{(1)}(w))^*$.

We therefore increment k to 1 and calculate

$$R_1^{(1)} = (-1-\delta)(1,1) + 3W_1$$

$|N_\delta(R_1^{(1)})| \approx 1.15$, $n = 1$ so we move to the calculation of a representation of $M_2^{(1)}(w)$ which gives

$$g_2^{(1)} = 1,\ \psi_2^{(1)} = 0,\ \sigma_2^{(1)} = \delta,\ \theta_2^{(1)} = (1-2\delta),\ h\,N_\delta(A_2^{(1)}) = 5$$

$$\text{(that is } W_2^{(1)} = ((1-2\delta)(1,1) + W)/5)$$

Since $N_\delta(R_1^{(1)})$ is not a root of unity we set

$$A_2^{(1)} = R_1^{(1)}A_1 = (-9-2\delta)(1,1) + (-6)W$$

and then note that

$$\lambda = 2g_2^{(1)}\,\theta_2^{(1)}/h\,N_\delta(A_2^{(1)}) = .4 - .8\delta \notin I_2'$$

which shows $M_2^{(1)}(w) \neq (M_2^{(1)}(w))^*$.

We now set $k = 2$ and calculate

$$R_2^{(1)} = \delta(1,1) + 2W_2$$

$|N_\delta(R_2^{(1)})| = .4$, $n = 1$. A representation of $M_3^{(1)}(w)$ is defined by

$$g_3^{(1)} = 2,\ \psi_3^{(1)} = 0,\ \sigma_3^{(1)} = \delta,\ \theta_3^{(1)} = 0,\ h\,N_\delta(A_3^{(1)}) = -2$$

$$\text{(that is } W_3^{(1)} = (0 + 2W)/(-2) = -W)$$

$N_\delta(R_2^{(1)})$ is not a root of unity so set

$$A_3^{(1)} = R_2^{(1)}A_2 = (-2-5\delta)(1,1) + (-6-2\delta)W$$

Since $\theta_3^{(1)} = 0$, $\psi_3^{(1)} = 0$ it is obvious that conditions (d), (e) of theorem 2.32 are satisfied and so we have $M_3^{(1)}(w) = (M_3^{(1)}(w))^*$. Thus the midpoint of a minimal period of length $q = 6$ has been found.

To complete the calculation of a symmetric minimal period we set

$$A_6 = A_3^{(1)}/(A_3^{(1)})^*, \quad A_4 = A_2^* A_6, \; A_5 = A_1^* A_6 \qquad //$$

This example illustrates two points which are typical of the cases with which we have dealt. The first point is that we generally find $n = 1$ at step 5 of algorithm 2.2. The second point is that we frequently find $\theta_{k+1} = \psi_{k+1} = 0$, $N_\delta(A_{k+1}) = \pm 2$ at the midpoint of an even length period. Thus the amount of calculation involved in recognizing the midpoint is indeed trivial in the cases which we have dealt with.

EXAMPLE 2.9

We now rework example 2.7b using algorithm 2.2. We have $\delta = \sqrt{-35}$, $\Upsilon = -3+5\delta$, $h = 1$, $w = \sqrt{\Upsilon/h}$. The initialization steps give $A_0 = (1,1)$ and

$$g_0 = 1, \; \psi_0 = 0, \; \sigma_0 = (1+\delta)/2, \; \theta_0 = 0, \; h N_\delta(A_0) = 1$$

The calculations for $k = 0$ give the following results.

$$R_0^{(1)} = ((15+3\delta)/2)(1,1) + 2W_0, \; n = 1, \; |N_\delta(R_0^{(1)})| \approx 18.1$$

$$g_1^{(1)} = 1, \; \psi_1^{(1)} = 0, \; \sigma_1^{(1)} = (1+\delta)/2, \; \theta_1^{(1)} = (31+5\delta)/2, \; hN_\delta(A_1^{(1)}) = (-21+5\delta)/2$$

$$A_1^{(1)} = ((15+3\delta)/2)(1,1) + 2W$$

$$\lambda = 2g_1^{(1)} \theta_1^{(1)}/hN_\delta(A_1^{(1)}) \approx .34 - .40\delta \notin I_1'$$

We therefore increment k to 1 where we find

$$R_1^{(1)} = ((-1+\delta)/2)(1,1) + 3W_1$$

$$R_1^{(2)} = (-3)(1,1) + ((1+\delta)/2)W_1$$

Thus n = 2. Since $|N_\delta(R_1^{(1)})| \approx .55$ we go to step 8. Since

$$N_\delta(R_1^{(2)}/R_1^{(1)}) \approx .94 - .33i$$

is clearly not a $Z(\delta)$ root of unity we discard $R_1^{(2)}$ and reset n = 1. $M_2^{(1)}(w)$ is defined by

$$g_2^{(1)} = 3,\ \psi_2^{(1)} = 1,\ \sigma_2^{(1)} = (5+\delta)/2,\ \theta_2^{(1)} = (-7+\delta)/2,\ h\,N_\delta(A_2^{(1)}) = (-19+\delta)/2$$

Since $N_\delta(R_1^{(1)})$ is not a root of unity we set

$$A_2^{(1)} = ((27-3\delta)/2)(1,1) + ((1-\delta)/2)W$$

and then note $M_2^{(1)}(w) \neq (M_2^{(1)}(w))^*$ since

$$\lambda = 2g_2^{(1)}\theta_2^{(1)}/h\,N_\delta(A_2^{(1)}) \approx 2.55 - .18\delta \notin I_2'$$

For k = 2 we again find n = 2.

$$R_2^{(1)} = (((-7-\delta)/2)(1,1) + 3W_2)/3$$

$$R_2^{(2)} = (((-7+\delta)/2)(1,1) + ((1-\delta)/2)W_2)/3$$

Since $|N_\delta(R_2^{(1)})| = 1$ we check to see if $N_\delta(R_2^{(j)})$ is a root of unity.

$$N_\delta(R_2^{(1)}) \approx .94 - .05\delta,\ N_\delta(R_2^{(2)}) = -1$$

Consequently steps 7, 8 of algorithm 2.2 effectively result in the discarding of $R_2^{(1)}$ plus the resetting of n = 1 and the resetting of

$$R_2^{(1)} = (((-7+\delta)/2)(1,1) + ((1-\delta)/2)W_2)/3$$

$M_3^{(1)}(w)$ is defined by

$$g_3^{(1)} = 3,\ \psi_3^{(1)} = 1,\ \sigma_3^{(1)} = (5+\delta)/2,\ \theta_3^{(1)} = (7-\delta)/2,\ h\,N_\delta(A_3^{(1)}) = (19-\delta)/2$$

Now $N_\delta(R_2^{(1)}) = -1$ so we must test to see if $M_3^{(1)}(w) = M_2^*(w)$. Since $\xi = -1$, $g_3^{(1)} = g_2$, $\psi_3^{(1)} = \psi_2$, $\sigma_3^{(1)} = \sigma_2$, $\theta_3^{(1)} = -\theta_2$ it is easily seen that

the conditions of theorem 2.32 are satisfied and so we do indeed have $M_3^{(1)}(w) = M_2^*(w)$. We have therefore found the midpoint of a minimal period of length $q = 5$. We therefore set

$$A_3^{(1)} = ((-61+13\delta)/2)(1,1) + ((3+3\delta)/2)W$$

and then complete the calculation of a symmetric minimal period by setting

$$A_5 = A_3^{(1)}/A_2^*,\ A_4 = A_1^* A_5 \qquad //$$

This example also illustrates points that are typical of the cases we have considered. Firstly the occurrence of $\theta_{k+1} = -\theta_k$ at the midpoint of a period of odd length is fairly common. Secondly if $n > 1$ in step 5 of algorithm 2.2 then we have always found that step 8 reduces n to 1. Finally we have generally found that for a given value of k either $|N_\delta(R_k^{(1)})| \neq 1$ or we are at the midpoint of an odd length period. We do not suggest that this will always be the case but we do note the simplicity and effectiveness of the test in step 6 of algorithm 2.2.

These examples are sufficient to illustrate the typical sort of calculations required and results obtained by algorithm 2.2. Note that in the next chapter we will make further use of algorithm 2.2 (and a modification of algorithm 2.2) to calculate units in certain types of quartic fields. (Of course this was the main reason for studying the subject of relative minima of $Z(\delta)$ modules). These unit calculations will therefore give further illustration of the performance of algorithm 2.2.

Apart from the completion of the proof of the bound for $|A_k^*|$ which was stated in section two we have now developed the general theory of

relative minima of M(w) sufficiently far for our purposes. (However in section two of the next chapter we will develop several specific results concerning modules of the form $M(\sqrt{\gamma})$ which are relevant to the problem of calculating units in quartic fields of the form $Q(\delta)(\sqrt{\gamma})$).

We therefore close this section with the promised proof of the bound for $|A_k^*|$.

THEOREM 2.33

Let $w \in \mathbb{C}\setminus\{0\}$ and let A_k, $k > 0$ be a relative minimum of M(w). Then

$$|A_k^*| < 2\sqrt{2}|\delta|/c\sqrt[4]{3}\pi^{\frac{1}{2}}(|\beta_k| - 1/2)$$

PROOF

Case (a) We begin with the special case where $w = \sqrt{\gamma}/h$. Since M(w) has periodic relative minima we can find A_q, $q > 0$ such that A_q/A_0 is a unit of M(w). Therefore by theorem 2.29b we have

$$A_{k+nq} = A_k(A_q/A_0)^n, \quad n \in Z$$

is also a relative minimum of M(w). Since A_q/A_0 is a unit we have $|N_\delta(A_{k+nq})| = |N_\delta(A_k)|$ and theorem 2.25 therefore gives

$$|N_\delta(A_k)| < 4\sqrt{2}|\delta||w|/c\sqrt[4]{3}\pi^{\frac{1}{2}} + r/|\beta_{k+nq}|$$

Now as $n \to \infty$ we have $|\beta_{k+nq}| \approx |A_{k+nq}|/2|w| \to \infty$. We therefore conclude

$$|N_\delta(A_k)| \leq 4\sqrt{2}|\delta||w|/c\sqrt[4]{3}\pi^{\frac{1}{2}} \tag{19}$$

and the result follows once we note that $|A_k^*| = |N_\delta(A_k)|/|A_k|$ and that

$$\begin{aligned} |A_k| &\geq 2|\beta_k w| - |A_k^*| \\ &= 2|w|(|\beta_k| - |A_k^*|/2|w|) \\ &> 2|w|(|\beta_k| - 1/2) \end{aligned}$$

since $|A_k^*| < |w|$ for $k > 0$.

Case (b) We now use case (a) to prove the general case where w is any non-zero complex number. Let $A_k = A_k^{(1)}, A_k^{(2)}, \ldots, A_k^{(n)}$ denote the complete set of elements of E_k. Throughout this proof we will use the following convention. If A denotes the element $\alpha(1,1) + \beta W \in M(w)$ then $A(x)$ for $x \in \mathbb{C}\backslash\{0\}$ will denote the element $\alpha(1,1) + \beta(x,-x) \in M(x)$. (Note that $A(w) = A$). We first prove a lemma which will be required later in the proof.

Lemma We can find $r \in \mathbb{R}^+$ such that for any $x \in \mathbb{C}\backslash\{0\}$ satisfying $|x-w| < r$ we have

$$\forall\ C \in M(w) \text{ such that } C \neq (0,0),\ C \neq A_k^{(j)},\ j = 1,2,\ldots,n$$

$$\text{either } |C(x)| > |A_k(x)| \quad \text{or} \quad |C^*(x)| > |A_k^*(x)|$$

(Note that since A_k is a relative minimum this result is trivially true for $x = w$. We want to prove that the result remains true for all x in some open set containing w, that is in all modules M(x) "close" to M(w)).

Proof Suppose the result is false. Then $\forall\ m \in Z^+$ we can find $C_m \in M(w)$, $x_m \in \mathbb{C}\backslash\{0\}$ such that

$$|x_m - w| < 1/m,\ |x_m| > |w|/2, C_m \neq (0,0),\ C_m \neq A_k^{(j)},\ j = 1,2,\ldots,n \tag{20}$$

$$|C_m(x_m)| \leq |A_k(x_m)|,\ |C_m^*(x_m)| \leq |A_k^*(x_m)| \tag{21}$$

Now $|A_k(x_m)| = |\alpha_k + \beta_k x_m|$, $|A_k^*(x_m)| = |\alpha_k - \beta_k x_m|$, $|x_m| < |w| + 1$ and so

$$|A_k(x_m)|,\ |A_k^*(x_m)| \leq |\alpha_k| + |\beta_k|(|w| + 1)$$

Therefore if $C_m = \alpha^{(m)}(1,1) + \beta^{(m)}W$ then we have

$$|\alpha^{(m)}| = |C_m(x_m) + C_m^*(x_m)|/2 < |\alpha_k| + |\beta_k|(|w| + 1)$$

$$|\beta^{(m)}| = |C_m(x_m) - C_m^*(x_m)|/2|x_m| < (|\alpha_k| + |\beta_k|(|w| + 1))/|x_m|$$

Since $|x_m| > |w|/2$ it follows that $|\alpha^{(m)}|$, $|\beta^{(m)}|$ are bounded independent of m. Now $Z(\delta)$ is discrete and it is therefore clear that there can only be finitely many distinct values of C_m. Consequently $C_m = C$ occurs for infinitely many m. Note that (20) implies $C \neq (0,0)$, $C \neq A_k^{(j)}$, $j = 1,2,\ldots,n$. However $x_m \to w$ as $m \to \infty$ and so (21) implies $|C| \leq |A_k|$, $|C^*| \leq |A_k^*|$. Since A_k is a relative minimum this implies either $C = (0,0)$ or $|C| = |A_k|$, $|C^*| = |A_k^*|$. That is either $C = (0,0)$ or $C = A_k^{(j)}$ for some $j \in \{1,2,\ldots,n\}$. This is a contradiction and so the lemma is proved.

Now to the proof of the general case. Fix $r \in \mathbb{R}^+$ for which the lemma applies. Note that the set of all numbers of the form $\sqrt{\gamma/h}$ is dense in $\mathbb{C}$. Therefore $\forall\ m \in Z^+$ we can choose $x_m \in \mathbb{C}\backslash\{0\}$ such that x_m is of the form $\sqrt{\gamma/h}$ and $|x_m - w| < r/m$.

For any particular value of m we can choose $A_k^{(\ell)}(x_m)$ with $\ell \in \{1,2,\ldots,n\}$ such that

$$|N_\delta(A_k^{(\ell)}(x_m))| \leq |N_\delta(A_k^{(j)}(x_m))|,\ j = 1,2,\ldots,n$$

In view of the manner in which $A_k^{(\ell)}(x_m)$ is chosen and the fact that the lemma applies it is easily checked that $A_k^{(\ell)}(x_m)$ is a relative minimum of $M(x_m)$. Since x_m is a number of the form $\sqrt{\gamma/h}$ it follows that (19) in case (a) applies to $A_k^{(\ell)}(x_m)$, that is

$$|N_\delta(A_k^{(\ell)}(x_m))| \leq 4\sqrt{2}|\delta||x_m|/c\sqrt[4]{3}\,\pi^{\frac{1}{2}}$$

Now consider what happens as $m \to \infty$. We have $x_m \to w$ and $|N_\delta(A_k^{(j)}(x_m))| \to |N_\delta(A_k^{(j)})| = |N_\delta(A_k)|$ for $j = 1,2,\ldots,n$. It is now easily seen that

$$|N_\delta(A_k)| \leq 4\sqrt{2}|\delta||w|/c\sqrt[4]{3}\,\pi^{\frac{1}{2}}$$

and the rest of the proof is the same as in case (a). //

SECTION FIVE

COMPARISONS AND CONCLUSIONS

In this section we briefly compare the work in the first four sections of this chapter with some of the existing relevant literature. In particular we note a generalization of the nearest integer continued fraction algorithm (algorithm 1.3) which applies in Euclidean complex quadratic fields, and an algorithm by Amara [1981] for calculating the class group and fundamental unit of a quadratic extension of a complex quadratic field for the case where the complex quadratic field has class number one. These two algorithms appear to be the only ones in the literature which deal specifically with some of the ideas which we have considered in this chapter. These comparisons will be extended in chapter three section two where we consider algorithms for calculating fundamental units of certain quartic fields. We then finish this section with several final comments concerning our work in this chapter.

The literature contains a number of works which consider generalizations of the simple continued fraction algorithm and related ideas. (Brentjes [1981] provides a survey of many of these works and gives a fairly extensive list of references to the relevant literature). The approaches used in generalizing the simple continued fraction algorithm fall into two main groups.

In the first group we find the algorithms which we can think of as generalizing the form of the simple continued fraction algorithm, that is successive approximations are calculated as a simple integral linear combination of previously obtained approximations. See for example the Jacobi-Perron algorithm (see Bernstein [1971]), and Szekeres [1970]. These algorithms tend to be fairly simple from a computational point

of view and have varied success in achieving the desired generalizations of the properties of the simple continued fraction algorithm. (As described say in Szekeres [1970,pp114,115]). In general these algorithms deal with real numbers and are therefore not particularly relevant to the work in this chapter. However there is a complex continued fraction algorithm which is relevant and we shall describe and briefly discuss it later in this section. (Szekeres has also indicated (verbally) that his algorithm can be applied in certain complex cases. However no details or results are yet available).

The second group of algorithms is based on the general idea of a relative minimum. (Usually defined in terms of the conjugates which correspond to some algebraic number field). Note the different approach in comparison to the first group. Here the objects of real interest, that is the relative minima, are defined first and an appropriate algorithm for their calculation is then developed. See for example Berwick [1913], Billevič [1956], and of course Voronoi [1896]. Many others have further developed the work of these authors, for example Rudman and Steiner [1978] (Berwick), Steiner and Rudman [1976] (Billevič), and Williams et al [1980] (Voronoi). Apart from Voronoi these algorithms tend to have a more restricted outlook in comparison to the first group. This is because they generally apply to Z modules whose basis elements form an integral basis of an algebraic number field and their sole purpose is to calculate a fundamental system of units for this field. Our algorithm therefore takes a slightly more general approach than much of the existing literature. (Our main reason for taking the more general approach was to highlight the close correspondence between the work in this chapter and the simple continued fraction algorithm. In the simple continued fraction case we approximate real numbers by using ratios of elements of the discrete set Z and in this chapter we

approximate complex numbers using ratios of elements of the discrete set $Z(\delta)$). In general we find that these relative minima algorithms are either specific to cubic fields (see especially Williams et al [1980]) or of a more general nature applying to arbitrary degree fields. Of course these general algorithms apply to integral bases of quartic fields of the form $Q(\delta)(\sqrt{\gamma/h})$ (the underlying field in section four) but nothing of a particularly specific nature such as the idea of minimal symmetric periods is developed. (Although Williams [1980] has considered a similar sort of idea in cubic fields). However there is one algorithm due to Amara [1981] which is specific to fields of the form $Q(\delta)(\sqrt{\gamma})$ where $\gamma \in Z(\delta)$ and $Q(\delta)$ has class number one. We shall discuss this algorithm immediately following the discussion of the previously mentioned complex continued fraction which begins in the next paragraph.

The ideas noted below concerning complex continued fractions have been dealt with to varying degrees by A. Hurwitz [1887], J. Hurwitz [1902], Arwin [1926], Stein [1927], and Lakein [1971,1974,1975] amongst others. Let

$$d = -1, -2, -3, -7, \text{ or } -11 \tag{22}$$

Then $Q(\delta)$ is Euclidean and so given any $w \in \mathbb{C}$ we can find $\varphi \in Z(\delta)$ such that $|\varphi - w| < 1$.

ALGORITHM 2.3

Let $w \in \mathbb{C}$ and assume d takes one of the values in (22). Then we can obtain the $Z(\delta)$CF ($Z(\delta)$ continued fraction) expansion

$$w = (\varphi_1, \varphi_2, \ldots), \varphi_k \in Z(\delta), \varphi_k \neq 0 \text{ for } k \geq 2$$

as follows.

1 Set $w_1 = w$, $k = 1$

2 Set $\varphi_k = \{w_k\}$

3 If $\varphi_k = w_k$ then stop

4 Set $w_{k+1} = (w_k - \varphi_k)^{-1}$

5 Increment k by 1, go to 2 //

Convergents ρ_k/τ_k ($\rho_k, \tau_k \in Z(\delta)$) are defined in the standard way and since $|w_k - \varphi_k| < 1$ it follows that the ρ_k/τ_k converge to w. Indeed most of the standard continued fraction results apply. (For example, $\rho_k\tau_{k-1} - \rho_{k-1}\tau_k = (-1)^k$, w has a finite $Z(\delta)$CF if and only if $w \in Q(\delta)$, etc.). Note that if $w \in \mathbb{R}$ then $\varphi_k \in Z$, that is the $Z(\delta)$CF for w and the nearest integer continued fraction (algorithm 1.3) for w are identical.

A $Z(\delta)$CF is said to be periodic if and only if $\exists\ m,r \in Z^+$ such that $w_m = \xi w_{m+r}$ where ξ is a $Z(\delta)$ root of unity. As usual m,r are assumed to be minimal and the terms preperiod, period have the obvious meaning. The expansion of a $Q(\delta)$ quadratic surd (that is an algebraic number of degree two over $Q(\delta)$) can be obtained using the obvious generalization of algorithm 1.2 and we have the result

$w \in \mathbb{C}$ has a periodic $Z(\delta)$CF

if and only if

w is a $Q(\delta)$ quadratic surd.

Consequently if w is a $Q(\delta)$ quadratic surd then we can obtain a unit of the quartic field $Q(\delta)(w)$ from the $Z(\delta)$CF expansion of w. We shall comment further on this point in chapter three section two.

The following points become evident when we compare the $Z(\delta)CF$ algorithm (algorithm 2.3 or the obvious generalization of algorithm 1.2) with our relative minima algorithm (algorithm 2.1 or 2.2). Firstly the two algorithms are not equivalent. In general we find that the convergents of the $Z(\delta)CF$ expansion of w correspond to a subsequence of a half chain of relative minima of M(w). Thus the $Z(\delta)CF$ algorithm tends to miss some of the "best approximations" to w. (A relative minimum is of course essentially a generalized best approximation). Of course this is not unexpected in view of the fact that the $Z(\delta)CF$ algorithm is a generalization of the nearest integer continued fraction algorithm rather than the simple continued fraction algorithm. (See the following paragraph for further details along these lines with respect to variations of the $Z(\delta)CF$ algorithm). A second point to note is that the $Z(\delta)CF$ algorithm is clearly much simpler from a computational point of view. However note that algorithms 2.1, 2.2 are also fairly simple as far as calculation is concerned when d takes one of the values in (22). This is because the search for R_k usually only involves testing 3 or 4 values of β at most (fewer for the smaller magnitude values of d), and the frequent occurrence of $\eta_{k,2} = 1$ means that the calculation of an $M_{k+1}(w)$ representation is generally fairly trivial. Thirdly we note that both algorithms can be used to find units of $Q(\delta)(w)$ when w is a $Q(\delta)$ quadratic surd and we shall expand on this point in chapter three. A final point to note is that our algorithm has one distinct advantage in that it is defined with respect to all complex quadratic fields rather than just those that are Euclidean.

We finish this discussion of complex continued fractions by noting that it is possible to define many variations of algorithm 2.3 by replacing step 2 with $\varphi_k = f(w_k)$ where

$$f : \mathbb{C} \to Z(\delta),\ |f(x)-x| < 1 \quad \forall\ x \in \mathbb{C}$$

Variations of this sort have been considered by J. Hurwitz [1902], Stein [1927], and Lakein [1971,1974,1975] and these algorithms have similar properties to the $Z(\delta)$CF algorithm. We also find that for a given $w \in \mathbb{C}$ it is often possible to choose f so that the convergents of the resulting continued fraction expansion of w agree with the relative minima of M(w). However examples such as $w = \sqrt{2\delta}$, $\delta = \sqrt{-11}$ show that this cannot always be done. In fact any complex continued fraction algorithm will fail to find infinitely many of the relative minima of M(w) for this particular value of w. In other words no complex continued fraction can be guaranteed to produce all (or even all but finitely many) of the best approximations to w. This is obviously a further point in favour of our algorithm.

This completes our discussion of complex continued fractions for the present. As has been noted at several points in this section we shall expand on this discussion with respect to calculation of units in chapter three section two.

Amara [1981] gives an algorithm for calculating the class number and fundamental unit of a quartic field of the form $Q(\delta)(\sqrt{\gamma})$ where $\gamma \in Z(\delta)$ and $Q(\delta)$ is a complex quadratic field having class number one, that is

$$d = -1, -2, -3, -7, -11, -19, -43, -67, \text{ or } -163$$

(See chapter three section one for more detail concerning this type of quartic field). We shall see that some of Amara's basic ideas are essentially the same as those involved in the development of algorithm 2.2. We begin by briefly describing Amara's work. For quartic fields of the form described above we have

$$Z(\delta)(\sqrt{Y}) = Z(\delta)[1,A],\ A \in Z(\delta)(\sqrt{Y}) \tag{23}$$

Let B'' denote the conjugate of $B \in Q(\delta)(\sqrt{Y})$ defined by the conjugate $(\sqrt{Y})'' = -\sqrt{Y}$ of $\sqrt{Y}$. A reduced ideal of $Z(\delta)(\sqrt{Y})$ is an ideal of the form

$$J = Z(\delta)[\tau, A-\mu],\quad \tau,\mu \in Z(\delta) \tag{24}$$

satisfying

$$\forall\ B \in J,\ |\tau| \leq \max\{|B|,|B''|\}$$

$Z(\delta)(\sqrt{Y})$ has finitely many reduced ideals ($|\tau|$ in (24) is bounded). With each reduced ideal Amara associates a number $H \in J$ referred to as an "elemént de conversion". The successor ideal of J is defined to be

$$s(J) = (H'/\tau)J$$

which is also a reduced ideal. For any reduced ideal there is a cycle of ideals

$$J_0, J_1, \ldots, J_n \text{ with } J_{k+1} = s(J_k) \text{ and } J_0 = s(J_n) \tag{25}$$

The finitely many reduced ideals of $Z(\delta)(\sqrt{Y})$ are therefore divided into a number of disjoint cycles and this number is the class number of $Q(\delta)(\sqrt{Y})$. Furthermore a fundamental unit of $Z(\delta)(\sqrt{Y})$ is given by

$$U = \frac{H_0}{\tau_0}\frac{H_1}{\tau_1}\cdots\frac{H_n}{\tau_n} \tag{26}$$

where H_k is the "elemént de conversion" of $J_k = Z(\delta)[\tau_k, A-\mu_k]$.

Amara's "elemént de conversion" is essentially a relative minimum of the corresponding ideal. The similarities between the above ideas and algorithm 2.2 are seen most clearly if we consider a case where we can take $A = \sqrt{Y}$ in (23). (Chapter three section one will show precisely when this occurs). If we take

$$J_0 = Z(\delta)(\sqrt{Y}) = Z(\delta)[1,\sqrt{Y}]$$

and compare the resulting cycle (25) with the results of applying algorithm 2.2 to $M(\sqrt{Y})$ we find $q = n + 1$,

$$J_k = Z(\delta)[N_\delta(A_{q-k}),\theta_{q-k} + \sqrt{Y}]$$

and up to a root of unity factor we have

$$H_k = \alpha N_\delta(A_{q-k}) + \beta(\theta_{q-k} + \sqrt{Y})$$

where

$$R_{q-k,q-k-1} = \alpha(1,1) + \beta W_k$$

(Recall that $g_k = 1$ when $h(d) = 1$). Finally up to a root of unity factor we have that the unit defined by (26) corresponds to A_q^*. Thus the differences are largely superficial, that is notation, a scale factor, and the fact that the cycle and the period run in opposite directions. The calculations required by the two approaches appear to be essentially the same although Amara does not take advantage of the symmetry which is present. (This would be particularly simple to do when $Q(\delta)$ is Euclidean since Amara shows that in such cases each reduced ideal has a unique standard form).

Thus there is a close connection between some of the basic ideas of Amara's work and our work with the main overlap in application of the two algorithms being in the calculation of units of certain quartic fields. (We shall comment further on this point in chapter three section two). Apart from this overlap the two algorithms have obviously differing aims.

We finish our comments on Amara's work by noting that Amara does not address the case where $h(d) > 1$. Consequently we can view our

work as being in some respects a generalization of Amara's work. More correctly we might suggest that our work in this chapter provides a possible avenue for generalizing Amara's work. However it is not our intention to pursue this final point in the present work.

We now finish this section and chapter with several concluding remarks concerning the work in this chapter. We note a number of areas which require further investigation (although we shall not do this in the present work) and then comment on how well we have succeeded in achieving the aims of this chapter as stated in the introductory paragraph.

There are several major points in this chapter which we have either been unable to resolve to our satisfaction or not considered in any depth. The first and most annoying of these points concerns the form of the midpoint of a period of even length. In theorem 2.31 we have only been able to prove that in such cases we will find

$$A_{k+1}, A_{k+1}^{(1)} \ (A_{k+1} \text{ arbitrary}) \text{ such that } M_{k+1}^{(1)}(w) = M_{k+1}^{*}(w)$$

However as noted in example 2.7a we have found in practice that we can always take $A_{k+1}^{(1)} = A_{k+1}$. We have been unable to resolve this point and further investigation will hopefully produce either the appropriate proof or a counter example. (Amara [1981, Lemme I.3] effectively proves the result for certain special cases when $h(d) = 1$). While this point is annoying from an aesthetic point of view it does not in fact cause any real problem from a practical point of view. This is because the amount of work required by algorithm 2.2 to take care of the possibility that in some cases we might not be able to take $A_{k+1}^{(1)} = A_{k+1}$ is relatively small. Of more importance from a practical point of view is the fact that the amount of work required by

algorithms 2.1A, 2.2A depends on $|\delta|$. Consequently for large $|\delta|$ the calculation of an R_k becomes more costly. The main reason for this dependence on $|\delta|$ is of course the fact that we have simply employed an exhaustive search technique to check through the possible β coefficients. It would therefore be appropriate to consider possible ways of streamlining this search especially for larger values of $|\delta|$. One question which we have not considered in this chapter is the possibility of generalizing (8) of chapter one. We have been unable to make any significant headway in this area and the development of an appropriate generalization (if indeed one exists) certainly deserves close attention. A final major point which deserves further consideration is the more general form of periodicity mentioned following theorem 2.23. We have not considered this point in any detail although it does appear that results of a more theoretical nature are relatively straight forward to develop. However there appear to be problems in developing an algorithm which has the same practical advantages that algorithm 2.2 has over algorithm 2.1.

In the introductory paragraph of this chapter we stated our objective was to generalize the ideas of chapter one sections four and three. We also indicated a more specific objective of developing an algorithm capable of calculating fundamental units of quadratic extensions of complex quadratic fields. In general we have been reasonably successful in developing appropriate generalizations of the ideas and results of chapter one sections four and three. The previous paragraph indicates the main areas where we have been less successful than we would have liked. On the final question of whether or not the work in this chapter has produced an algorithm capable of calculating fundamental units of quadratic extensions of complex quadratic fields we simply note that the answering of this question forms a major part of the following chapter.

CHAPTER THREE

UNITS OF CERTAIN QUARTIC EXTENSIONS OF Q HAVING A QUADRATIC SUBFIELD

Quartic extensions of the rational field can be divided into two groups according to the presence or absence of a quadratic subfield. In this chapter we consider only those quartic extensions of Q which have a quadratic subfield. In section one we look at results concerning the integers of such fields. In particular we express the integers of the quartic extension in terms of the integers of the quadratic subfield. We also look at the unit group structure of these quartic fields and classify them accordingly into four types. In sections two and three we shall develop algorithms for calculating a fundamental system of units in two of these four types of field. These algorithms rely heavily on the work of chapter two, in particular section four.

SECTION ONE

QUARTIC FIELDS HAVING A QUADRATIC SUBFIELD, THEIR INTEGERS AND A CLASSIFICATION

Throughout this chapter we assume that $Q(\Gamma)$ is a quartic extension of Q which has a quadratic subfield $Q(\delta)$ (see chapter one section two).

DEFINITION 3.1

Let $\gamma \in Z(\delta)$. We say that

γ is non-square and rational square-free

if and only if

$\sqrt{\gamma} \notin Z(\delta)$ and γ is not divisible by the square of a rational prime.

Such a γ can be written as

$$\gamma = a_1 a_2 \rho, \quad a_1, a_2 \in Z^+, \quad \rho \in Z(\delta)$$

where $a_1 a_2$ is square-free, $(a_1, \Delta) = 1$, $a_2 | \Delta$ (Δ is the discriminant of $Q(\delta)$), and ρ is not divisible by a rational prime. //

The following result shows that $Q(\Gamma)$ can be viewed as a quadratic extension of $Q(\delta)$. (See also Nagell [1961, Theorem 2]).

THEOREM 3.1

Let $Q(\Gamma)$ be a quartic extension of Q which has a quadratic subfield. Then

$$Q(\Gamma) = Q(\delta)(\sqrt{\gamma}) = Q(\delta + \sqrt{\gamma})$$

where γ is a non-square rational square-free $Z(\delta)$ integer. Furthermore if $\gamma \notin Z$ then

$$Q(\Gamma) = Q(\sqrt{\gamma})$$

PROOF

We have $[Q(\Gamma) : Q] = [Q(\Gamma) : Q(\delta)][Q(\delta) : Q] = 4$. Clearly Γ is of degree 2 over $Q(\delta)$. We can therefore find $\kappa, \lambda \in Q(\delta)$ such that

$$\Gamma^2 + 2\kappa\Gamma + \lambda = 0$$

which implies

$$\Gamma = -\kappa + s\sqrt{\kappa^2 - \lambda}, \qquad s \in \{-1, 1\}$$

$$= -\kappa + \theta\sqrt{\gamma}, \qquad \gamma \in Z(\delta), \ \theta \in Q(\delta)$$

Thus $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ and we can clearly choose γ to be non-square and rational square-free.

To show that $Q(\Gamma) = Q(\delta+\sqrt{\gamma})$ we need only show that $\delta + \sqrt{\gamma}$ is not of degree 1 or 2 over Q. If $\delta + \sqrt{\gamma}$ was of degree 1 or 2 over Q then we would have $\delta + \sqrt{\gamma} = a + b\sqrt{e}$ for some $a,b,e \in Q$. However this would then imply $(\delta+\sqrt{\gamma}-a)^2 = b^2 e$ which when expanded implies $\sqrt{\gamma} \in Q(\delta)$. This is a contradiction and so we conclude that $\delta + \sqrt{\gamma}$ is of degree 4 over Q.

Finally for $\gamma \notin Z$ we have $Q(\delta) = Q(\gamma)$ and so

$$Q(\delta)(\sqrt{\gamma}) = Q(\gamma)(\sqrt{\gamma}) = Q(\sqrt{\gamma}) \qquad //$$

Note that we could have required γ in theorem 3.1 to be square-free. We have not done so for two reasons. Firstly, although rational square divisors of γ are easily determined it is often quite difficult to determine non-rational square factors of γ when the class number of $Q(\delta)$ satisfies $h(d) > 1$. Secondly, insisting that γ is square-free does not necessarily imply that $Q(\Gamma)$ has a unique representation of the form $Q(\delta)(\sqrt{\gamma})$. This is because for $h(d) > 1$ it is possible to have $Q(\delta)(\sqrt{\gamma}) = Q(\delta)(\sqrt{\kappa})$ with γ,κ distinct square-free $Z(\delta)$ integers. For example, if $\delta = \sqrt{-10}$ then

$$(3+2\delta)(5+\delta)^2 = 5(3+2\delta)^2, \quad 5(2)^2 = -2(\delta)^2$$

and so we have

$$Q(\delta)(\sqrt{3+2\delta}) = Q(\delta)(\sqrt{5}) = Q(\delta)(\sqrt{-2})$$

It is therefore not always immediately obvious whether or not two fields are distinct. However it is not difficult to deduce that $Q(\delta)(\sqrt{\gamma}) = Q(\delta)(\sqrt{\kappa})$ if and only if $\kappa\gamma = \alpha^2$, $\alpha \in Z(\delta)$, $\alpha \neq 0$.

Viewing $Q(\Gamma)$ as a quadratic extension of $Q(\delta)$ proves most useful in developing further results. Some of the results developed using this approach correspond closely with results from the standard quadratic case.

In the first such result we show that the integers of $Q(\delta)(\sqrt{\gamma})$ can be expressed in terms of the integers of $Q(\delta)$. (In some cases more specific results can be found in the literature. For example, see Hilbert [1894] for the case $Q(i)(\sqrt{\gamma})$ and Williams [1970] for the case $Q(\delta)(\sqrt{d_1}), d_1 \in Z$).

THEOREM 3.2

Suppose $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ where γ is a non-square rational square-free $Z(\delta)$ integer. Let $\gamma = \gamma_1\gamma_2$ where $\gamma_1, \gamma_2 \in Z(\delta)$ and γ_1 satisfies

$$\forall\ \beta \in Z(\delta),\ \gamma_1 | \beta^2 \text{ implies } \gamma_1 | \beta$$

(For example, we can take $\gamma_1 = 1$, $\gamma_2 = \gamma$. However we shall see that other more desirable possibilities often exist). Then the ring of integers of $Q(\delta)(\sqrt{\gamma})$ is

$$Z(\delta)(\sqrt{\gamma}) = \{ (\alpha+(\beta/\gamma_2)\sqrt{\gamma})/2 : \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2),\ \alpha^2 \equiv \beta^2 \gamma_1/\gamma_2 \pmod 4 \}$$

where

$$I(\gamma_1,\gamma_2) = \{\beta \in Z(\delta) : \gamma_2 | \beta^2\gamma_1\}$$

is an ideal of $Z(\delta)$.

PROOF

Suppose $A \in Z(\delta)(\sqrt{\gamma})$. Since $A \in Q(\delta)(\sqrt{\gamma})$ we have $A = \theta + \lambda\sqrt{\gamma}$ with $\theta,\lambda \in Q(\delta)$. The conjugate of A with respect to the quadratic extension $Q(\Gamma)$ of $Q(\delta)$ is $A^* = \theta - \lambda\sqrt{\gamma}$. (Note that there is no real conflict between this usage of A^* and the usage of A^* in chapter two). Since $A, A^*, \sqrt{\gamma}$ are algebraic integers it follows that

$$A + A^* = 2\theta,\ \sqrt{\gamma}(A-A^*) = 2\lambda\gamma,\ AA^* = \theta^2 - \lambda^2\gamma$$

are $Z(\delta)$ integers. Let $2\theta = \alpha$, $2\lambda\gamma = \kappa$ where $\alpha,\kappa \in Z(\delta)$. Now $AA^* = (\alpha^2-\kappa^2/\gamma)/4 \in Z(\delta)$ implies $\gamma | \kappa^2$. Consequently $\gamma_1 | \kappa$, that is $\kappa = \gamma_1\beta$ with $\beta \in Z(\delta)$. Since $\gamma | \kappa^2$ is equivalent to $\gamma_2 | \beta^2\gamma_1$ we have

$\beta \in I(\gamma_1,\gamma_2)$. Thus

$$A = (\alpha+(\beta/\gamma_2)\sqrt{\gamma})/2,\ \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2) \tag{1}$$

and the condition $AA^* = (\alpha^2-\beta^2\gamma_1/\gamma_2)/4 \in Z(\delta)$ is equivalent to

$$\alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4 \tag{2}$$

Now suppose that A is of the form in (1) and α,β satisfy the conditions in (1) and (2). Then A is a root of the monic polynomial

$$f(x) = (x-A)(x-A^*) = x^2 - \alpha x + (\alpha^2-\beta^2\gamma_1/\gamma_2)/4$$

The conditions in (1), (2) ensure that the coefficients of $f(x)$ are $Z(\delta)$ integers and so A is an algebraic integer. Since $A \in Q(\delta)(\sqrt{\gamma})$ it follows that $A \in Z(\delta)(\sqrt{\gamma})$.

It remains to show that $I(\gamma_1,\gamma_2)$ is an ideal. Suppose $\beta,\lambda \in I(\gamma_1,\gamma_2)$. Then $(\beta/\gamma_2)\sqrt{\gamma}$, $(\lambda/\gamma_2)\sqrt{\gamma} \in Z(\delta)(\sqrt{\gamma})$. Consequently $((\beta+\lambda)/\gamma_2)\sqrt{\gamma} \in Z(\delta)(\sqrt{\gamma})$ which implies $\beta+\lambda \in I(\gamma_1,\gamma_2)$. The remaining conditions which $I(\gamma_1,\gamma_2)$ must satisfy in order to be an ideal are trivial to check. //

This result does not of course have the simplicity of the corresponding result in the standard quadratic case. Furthermore in view of the above theorem it is not surprising to find that $Z(\delta)(\sqrt{\gamma})$ cannot always be expressed as a $Z(\delta)$ module. For example MacKenzie and Scheuneman [1971] have shown that $Z(\sqrt{-14})(\sqrt{-7})$ cannot be expressed as $Z(\sqrt{-14})$ module. In fact such occurrences are quite common when $h(d) > 1$. (Note that this problem is not a consequence of our not insisting that γ be square-free). Clearly a more precise description of $Z(\delta)(\sqrt{\gamma})$ is not going to be an easy matter to determine.

Before looking more closely at the general form of $Z(\delta)(\sqrt{\Upsilon})$ we illustrate theorem 3.2 with an example.

EXAMPLE 3.1

Let $\delta = \sqrt{10}$, $\Upsilon = 12 + 3\delta = 3(4+\delta)$. Υ is clearly rational square-free and it is easily checked that Υ is non-square. (In fact we have $N(\Upsilon) = 54 = 2.3^3$ and since $N(\alpha) = 3$ cannot occur for $\alpha \in Z(\delta)$ we see that Υ is actually square-free). $Z(\delta)(\sqrt{\Upsilon})$ is determined as follows.

It is a simple matter to check that $\forall\ \beta \in Z(\delta)$, $3|\beta^2$ implies $3|\beta$. We can therefore take $\Upsilon_1 = 3$, $\Upsilon_2 = 4 + \delta$. The implications of congruence (2) are now determined. First note that $\Upsilon_1^2 \equiv 1 \pmod 4$. Therefore

$$\alpha^2 \equiv \beta^2\Upsilon_1/\Upsilon_2 \pmod 4 \text{ implies } \Upsilon_1\Upsilon_2\alpha^2 \equiv \beta^2 \pmod 4$$

Now if $\kappa \in Z(\delta)$ then $\kappa \equiv 0,1,\delta,1+\delta \pmod 2$. Thus

$$\kappa^2 \equiv 0, 1, 2, 3+2\delta \pmod 4$$

$$\Upsilon_1\Upsilon_2\kappa^2 \equiv 0, 3\delta, 2\delta, \delta \pmod 4$$

Clearly $\Upsilon_1\Upsilon_2\alpha^2 \equiv \beta^2 \pmod 4$ requires $\alpha \equiv \beta \equiv 0 \pmod 2$. It is now not difficult to deduce that

$$Z(\delta)(\sqrt{\Upsilon}) = \{\alpha+(\beta/\Upsilon_2)\sqrt{\Upsilon} : \alpha \in Z(\delta),\ \beta \in I(\Upsilon_1,\Upsilon_2)\} \qquad (3)$$

We now need to determine $I(\Upsilon_1,\Upsilon_2)$. Obviously $\Upsilon_2 \nmid (1)^2\Upsilon_1$ and it is easily checked that $\Upsilon_2|(2)^2\Upsilon_1, \Upsilon_2|(\delta)^2\Upsilon_1$. It therefore follows that we have $I(\Upsilon_1,\Upsilon_2) = Z[2,\delta]$. Note that

$$I(\Upsilon_1,\Upsilon_2) \supsetneq <\Upsilon_2> = Z[6,4+\delta]$$

We can now conclude

$$Z(\delta)(\sqrt{\Upsilon}) = \{\alpha+(\beta/\Upsilon_2)\sqrt{\Upsilon} : \alpha \in Z(\delta),\ \beta \in Z[2,\delta]\}$$

Note that $Z(\delta)(\sqrt{\gamma})$ is not a $Z(\delta)$ module. If it was then we would have $Z(\delta)(\sqrt{\gamma}) = Z(\delta)[A,B]$ with

$$A = \alpha + (\beta/\gamma_2)\sqrt{\gamma},\ B = \kappa + (\lambda/\gamma_2)\sqrt{\gamma} \in Z(\delta)(\sqrt{\gamma})$$

Since $1 \in Z(\delta)(\sqrt{\gamma})$ there must exist $\theta, \eta \in Z(\delta)$ such that $\theta A + \eta B = 1$, that is $\theta\alpha + \eta\kappa = 1$. It therefore follows that

$$Z(\delta)(\sqrt{\gamma}) = Z(\delta)[\theta A + \eta B, -\kappa A + \alpha B]$$

$$= Z(\delta)[1, ((\alpha\lambda - \kappa\beta)/\gamma_2)\sqrt{\gamma}]$$

This clearly implies $<\alpha\lambda - \kappa\beta> = I(\gamma_1,\gamma_2) = Z[2,\delta]$ which is a contradiction since $Z[2,\delta]$ is not principal. //

In general $Z(\delta)(\sqrt{\gamma})$ is not as easily determined as in the above example. The main difficulty lies in determining $I(\gamma_1,\gamma_2)$. Exhaustive testing for a basis of $I(\gamma_1,\gamma_2)$ is obviously a totally impractical procedure in general and so a more efficient alternative is required. Checking γ_1 possibilities also presents some difficulty in general. We could of course simply avoid this problem by taking $\gamma_1 = 1$, $\gamma_2 = \gamma$. However this choice is not always desirable when other options exist. In particular if we can take $\gamma_1 = \gamma$ then the form of $Z(\delta)(\sqrt{\gamma})$ is considerably simplified. (See theorem 3.4 below). Finally congruence (2) also presents some difficulty in that it often doesn't lead to a simple single condition such as $\alpha \equiv \beta \equiv 0 \pmod 2$ or $\alpha \equiv \beta \pmod 2$. However it will prove useful to know exactly when (2) implies $\alpha \equiv \beta \equiv 0 \pmod 2$ since this leads to the simplified form of $Z(\delta)(\sqrt{\gamma})$ given in (3). In the following paragraphs we take a closer look at the points mentioned in this paragraph.

We begin by looking at a way of checking a given γ_1 possibility. That is we look at a way of determining the truth value of the statement

$$\forall\ \beta \in Z(\delta),\ \gamma_1 | \beta^2 \ \text{ implies } \ \gamma_1 | \beta \tag{4}$$

If $h(d) = 1$ then it is easily seen that (4) is true if and only if γ_1 is square-free. However if $h(d) > 1$ holds then (4) is not necessarily true for γ_1 square-free. The following theorem gives a precise description of those γ_1 for which (4) is true.

THEOREM 3.3

Let $\gamma_1 \in Z(\delta)$. Then the following statements are equivalent.

(a) $\forall\ \beta \in Z(\delta),\ \gamma_1 | \beta^2$ implies $\gamma_1 | \beta$

(b) $<\gamma_1> = P_1 P_2 \ldots P_m$, P_j distinct prime ideals of $Z(\delta)$

(c) $|N(\gamma_1)| = p_1^{a_1} p_2^{a_2} \ldots p_n^{a_n}$, p_j distinct rational primes, $a_j \in \{1,2\}$ with $a_j = 2$ if and only if $p_j | \gamma_1$ and $p_j \nmid \Delta$

PROOF

If γ_1 is a unit then (a),(b), (c) are trivially true. Therefore assume γ_1 is not a unit.

To prove (a) and (b) are equivalent it clearly suffices to prove that

(a1) $\forall$ principal ideals $I \subseteq Z(\delta)$, $<\gamma_1> | I^2$ implies $<\gamma_1> | I$

is equivalent to (b). Now the ideals of $Z(\delta)$ factor uniquely and so it is trivial to prove (b) implies (a1). We prove (a1) implies (b) by showing that whenever (b) is false we can construct an ideal I such that $<\gamma_1> | I^2$ but $<\gamma_1> \nmid I$.

If (b) is false then we can write $<\gamma_1> = P_1^{b_1} P_2^{b_2} \ldots P_m^{b_m}$, P_j distinct prime ideals, $b_j \in Z^+$, $b_1 \geq 2$. Let

$$H = (P_2^{b_2} \ldots P_m^{b_m})^{h(d)}$$

Note that H is principal, $P_1 \nmid H$, $<\gamma_1> | P_1^{b_1} H$. Now the ideal P_1 is

either principal, or not self-conjugate, or self-conjugate and not principal. (See theorem 1.3).

If P_1 is principal then take $I = P_1^{b_1-1} H$. Clearly I is principal and $<\gamma_1> \nmid I$. However since $b_1 \geq 2$ we have $<\gamma_1> | I^2$.

If P_1 is not self-conjugate then take $I = P_1^{b_1-1} (P_1')^{b_1-1} H$. ($P_1'$ is the conjugate ideal of P_1). Now $P_1 P_1'$ is principal and so I is also principal. Furthermore $P_1 \neq P_1'$ implies $<\gamma_1> \nmid I$. However we clearly have $<\gamma_1> | I^2$.

If P_1 is self conjugate but not principal then the choice of I is more involved. From theorem 1.3 we can deduce that $P_1^2 = <p>$, p a rational prime, $p|\Delta$, and

$$P_1 = \begin{cases} <2, 1+\delta> & \text{if } p = 2,\ d \equiv 3 \pmod 4 \\ <p, \delta> & \text{otherwise} \end{cases}$$

Now let

$$P = \begin{cases} <(d-1)/2, 1+\delta> & \text{if } p = 2,\ d \equiv 3 \pmod 4 \\ <d/p, \delta> & \text{otherwise} \end{cases}$$

Note that in the otherwise case p does indeed divide d since $p|\Delta$, $p \nmid d$ can only occur when $p = 2$ and $d \equiv 3 \pmod 4$. A simple calculation now shows that

$$P_1 P = \begin{cases} <1+\delta> & \text{if } p = 2,\ d \equiv 3 \pmod 4 \\ <\delta> & \text{otherwise} \end{cases}$$

Thus PP_1 is principal. Furthermore $(2,(d-1)/2) = 1$ when $d \equiv 3 \pmod 4$, and in the otherwise case $(d/p,p) = 1$ since d is square-free. Thus $P_1 \neq P$, or equivalently $P_1 \nmid P$. Now let $I = P_1^{b_1-1} P^{b_1-1} H$. Then as

before we have I principal, $<\gamma_1>|I^2$ but $<\gamma_1>\nmid I$. This completes the proof of the equivalence of (a) and (b).

We begin the proof of the equivalence of (b) and (c) by showing that (b) implies (c). From the description of the prime ideals of $Z(\delta)$ given in theorem 1.3 it is easily deduced that if $<\gamma_1>$ is the product of distinct prime ideals then we can rewrite the product in (b) as

$$<\gamma_1> = P_1P_2\ldots P_k(P_{k+1}P'_{k+1})(P_{k+2}P'_{k+2})\ldots(P_nP'_n)$$

where $(N(P_j), N(P_\ell)) = 1$ for $j \neq \ell$. $(P_{k+1}, P_{k+2}, \ldots, P_n$ are clearly type (ii) ideals. (See theorem 1.3 for type (i), (ii), and (iii) ideals)). Now $N(P_j) = p_j^2$ if P_j is a type (i) ideal and $N(P_j) = p_j$ otherwise. (p_j is some rational prime). Therefore

$$\begin{aligned} |N(\gamma_1)| &= N(<\gamma_1>) \\ &= N(P_1)\ldots N(P_k)\, N(P_{k+1})^2 \ldots N(P_n)^2 \\ &= p_1^{a_1}\ldots p_k^{a_k}\, p_{k+1}^2 \ldots p_n^2 \end{aligned}$$

where the p_j are distinct rational primes and $a_j \in \{1,2\}$. Now if $j \geq k+1$ then $P_jP'_j = <p_j>|<\gamma_1>$ and $(\Delta/p_j) = 1$, that is $p_j|\gamma_1$, $p_j\nmid\Delta$. If $a_j = 2$ then P_j must be a type (i) ideal and so we have $P_j = <p_j>|<\gamma_1>$, $(\Delta/p_j) = -1$, that is $p_j|\gamma_1, p_j\nmid\Delta$. Finally if $a_j = 1$ it is easily checked that $p_j\nmid\gamma_1$. Thus $|N(\gamma_1)|$ satisfies the conditions in (c).

To complete the proof of the theorem we show that (b) false implies (c) false. Therefore suppose $P^2|<\gamma_1>$, P a prime ideal. If P is a type (i) ideal then clearly $N(P^2) = p^4|N(\gamma_1)$ and so (c) is false. If P is a type (ii) ideal then $N(P^2) = p^2|N(\gamma_1)$. Now if $p\nmid\gamma_1$ then (c) is false. However if $p|\gamma_1$ then we also have that (c) is false. This

is because $p|\gamma_1$ implies $<p> = PP'|<\gamma_1>$ which implies $P^2P'|<\gamma_1>$ and so $N(P^2P') = p^3|N(\gamma_1)$. Finally if P is a type (iii) ideal then $P^2 = <p>|<\gamma_1>$ and $(\Delta/p) = 0$. Thus $p^2|N(\gamma_1)$, $p|\gamma_1$, $p|\Delta$ and so (c) is false. //

Thus to check if a given γ_1 satisfies (4) we need only look at the factorization of $N(\gamma_1)$. Of particular interest when determining the form of $Z(\delta)(\sqrt{\gamma})$ is the possibility of choosing $\gamma_1 = \gamma$.

THEOREM 3.4

Let γ be a non-square rational square-free $Z(\delta)$ integer and suppose that $N(\gamma)$ satisfies condition (c) of theorem 3.3. Then we can choose $\gamma_1 = \gamma$, $\gamma_2 = 1$. Consequently

$$I(\gamma_1,\gamma_2) = I(\gamma,1) = Z(\delta)$$

and so

$$Z(\delta)(\sqrt{\gamma}) = \{(\alpha+\beta\sqrt{\gamma})/2 : \alpha,\beta \in Z(\delta),\ \alpha^2 \equiv \beta^2\gamma \pmod 4\}$$

PROOF

The theorem follows immediately from theorems 3.2, 3.3. //

Note that if $h(d) = 1$ then we can always express the integers of $Q(\Gamma)$ is this simple form. To do this we choose γ square-free rather than just rational square-free. (This is relatively easy to do when $h(d) = 1$). Then as noted prior to theorem 3.3 we have that $\gamma_1 = \gamma$ satisfies (4) and so the result of theorem 3.4 applies.

More generally recall that we have $\gamma = a_1a_2\rho$. (See definition 3.1). Now theorem 3.3 shows that we cannot have $a_2|\gamma_1$ if $a_2 > 1$. Consequently if $a_2 > 1$ we can try the possibility $\gamma_1 = \gamma/a_2 = a_1\rho$ rather than $\gamma_1 = \gamma$. If both of these possibilities fail then unless factors of ρ are easily determined we will normally settle for $\gamma_1 = a_1$.

(Note that theorem 3.3 shows that (4) is always satisfied for this choice of γ_1).

Theorem 3.4 shows that the determination of $I(\gamma_1,\gamma_2)$ is trivial when we can take $\gamma_1 = \gamma$. In the remaining cases the next theorem indicates a relatively simple method for calculating $I(\gamma_1,\gamma_2)$.

THEOREM 3.5

Let $\gamma = a_1 a_2 \rho = \gamma_1\gamma_2$ be as in definition 3.1 and theorem 3.2. Then

$$I(1,\gamma) = Z[h,\gamma,\omega\gamma,\alpha]$$

where

$$h = p_1^{n_1} p_2^{n_2} \dots p_r^{n_r}, \quad n_j = [(m_j+1)/2]$$

$$a_1 a_2 N(\rho) = p_1^{m_1} p_2^{m_2} \dots p_r^{m_r}$$

$$\alpha = a_1\rho(d+\delta)$$

Furthermore

$$I(1,\gamma) = <\gamma_1> I(\gamma_1,\gamma_2)$$

and so

$$I(\gamma_1,\gamma_2) = Z[h/\gamma_1,\gamma_2,\omega\gamma_2,\alpha/\gamma_1]$$

PROOF

We begin by showing that

$$S = Z[h,\gamma,\omega\gamma,\alpha] \subseteq I(1,\gamma)$$

Clearly $\gamma,\omega\gamma \in I(1,\gamma)$. Since $\gamma\rho' = a_1 a_2 N(\rho) | h^2$ it is also clear that $h \in I(1,\gamma)$. In fact since $\gamma | k^2$, $k \in Z$ requires $a_1 a_2 N(\rho) | k^2$ it is not difficult to see that h is the smallest positive rational integer in $I(1,\gamma)$. It therefore remains to show that $\alpha \in I(1,\gamma)$. We have

$$\alpha^2 = \gamma a_1\rho(d(d+1) + 2d\delta)/a_2$$

Now a_2 is square-free and $a_2|\Delta$. It therefore follows that $a_2|d(d+1),2d$. Thus $(d(d+1) + 2d\delta)/a_2 \in Z(\delta)$ and so $\gamma|\alpha^2$, that is $\alpha \in I(1,\gamma)$.

To show that $S = I(1,\gamma)$ we must show that S contains a basis of $I(1,\gamma)$. Note that the minimal nature of h implies

$$I(1,\gamma) = Z[h,\theta],\ \theta = (e+f\delta)/c \in Z(\delta)$$

where f is the minimal positive coefficient of δ occurring in $I(1,\gamma)$ and $0 \leq e < hc$. (See theorem 1.4). We therefore need to show that $\theta \in S$. Let $f_1 = a_1(a_2,N(\rho))$. We shall show that $f_1|f$ and then show that S contains an integer of the form $\theta_1 = (e_1+f_1\delta)/c$. Since $S \subseteq I(1,\gamma)$ this implies $f_1 = f$ and it is then not difficult to check that $\theta \in S$.

Since $\gamma|\theta^2$ we have $a_1|\theta$. (See the comments following theorem 3.4). Thus $a_1|f$. Now a_2 is square-free and so

$$(a_2,N(\rho)) = p_1p_2\ldots p_r,\ p_j \text{ distinct rational primes}$$

Clearly $p_j|\Delta$. This implies $<p_j> = P_j^2$, P_j a self conjugate prime ideal. Since $p_j|N(\rho)$ we also have $P_j|<\rho>$. We therefore have $P_j^3|<\gamma>$. Thus $<\gamma>|<\theta>^2$ implies $P_j^2|<\theta>$, that is $p_j|\theta$ and so $p_j|f$. Consequently $(a_2,N(\rho))|f$. Since $(a_1,a_2) = 1$ we can now conclude that $f_1|f$.

We now show that S contains an integer of the form $\theta_1 = (e_1+f_1\delta)/c$. From theorem 1.5 we see that $<\gamma> = Z[\gamma,\omega\gamma]$ contains an integer of the form $\theta_2 = (a+a_1a_2\delta)/c$. Clearly $<\gamma> \subseteq S$ and so $\theta_2 \in S$. The required θ_1 will prove to be a linear combination of θ_2 and α. Note that the argument in the previous paragraph can also be used to prove $f_1|\beta,\ \forall\ \beta \in I(1,\gamma)$. In particular $f_1|\alpha$ and so

$$\alpha = a_1\rho(d+\delta) = f_1(k_1+k_2\delta)/c,\ (k_1+k_2\delta)/c \in Z(\delta)$$

To prove that the required θ_1 is a linear combination of α and θ_2 it suffices to prove that $(a_1a_2, f_1k_2) = f_1$ or equivalently $(a_1a_2/f_1, k_2) = 1$. (Note that $\theta_2 \in I(1,\gamma)$ and so $f_1 | \theta_2$ which implies $f_1 | a_1a_2$). Therefore with the intention of producing a contradiction we assume $p | (a_1a_2/f_1, k_2)$, p a rational prime. It is easily checked that $p | a_2$, $p | \Delta$, $p \nmid f_1$, $p \nmid N(\rho)$. Furthermore $p | \Delta$ implies $\langle p \rangle = P^2$ where P is a self-conjugate prime ideal. Now if $p \nmid d$ then we must have $p = 2$, $d \equiv 3 \pmod 4$ and so $p | (d-1)$. Thus $p | N(d+\delta) = d(d-1)$ and this implies $p | N(\alpha)$. Since $p \nmid f_1$ we have $p | N(\alpha/f_1) = (k_1^2 - k_2^2 d)/c^2$. Now we are assuming $p | k_2$ and it therefore follows that $p | k_1$. Since $c = p = 2$ cannot occur ($c = 2$ implies Δ is odd, and $p | \Delta$) we have $p | (k_1+k_2\delta)/c$, that is $p | \alpha$. In terms of ideals we therefore have $P^2 | \langle \alpha \rangle$. Now $p \nmid a_1$, $p \nmid N(\rho)$ implies $P \nmid \langle a_1\rho \rangle$ and so $P^2 | \langle d+\delta \rangle$, that is $p | (d+\delta)$. This is clearly a contradiction and so the proof of $S = I(1,\gamma)$ is complete.

The remaining results in the theorem are trivial to prove. //

Thus $I(\gamma_1,\gamma_2)$ can be determined fairly easily by first calculating h and then reducing the module $Z[h/\gamma_1, \gamma_2, \omega\gamma_2, \alpha/\gamma_1]$ to the standard representation $Z[g,\sigma]$ using the technique illustrated in example 1.1.

We now consider the congruence $\alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4$, $\alpha \in Z(\delta)$, $\beta \in I(\gamma_1,\gamma_2)$. As was noted earlier in this section we shall confine our attention to determining when this congruence implies $\alpha \equiv \beta \equiv 0 \pmod 2$.

THEOREM 3.6

Let $\gamma, \gamma_1, \gamma_2$ be as in theorem 3.2 and suppose $Z[g,\sigma]$ is the standard representation of $I(\gamma_1,\gamma_2)$. Then

$$\forall\, \alpha \in Z(\delta), \forall \beta \in I(\gamma_1,\gamma_2), \alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4 \text{ implies } \alpha \equiv \beta \equiv 0 \pmod 2 \qquad (5)$$

if and only if

$$\alpha^2 \not\equiv g^2\gamma_1/\gamma_2 \pmod 4,\ \alpha^2 \not\equiv \sigma^2\gamma_1/\gamma_2 \pmod 4,\ \alpha^2 \not\equiv (g+\sigma)^2\gamma_1/\gamma_2 \pmod 4 \qquad (6)$$

for $\alpha \in \{0,1,\omega,1+\omega\}$

When (5) is true we have

$$Z(\delta)(\sqrt{\gamma}) = \{\alpha+(\beta/\gamma_2)\sqrt{\gamma} : \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2)\}$$

PROOF

Suppose (5) is true. Then given any $\alpha \in Z(\delta)$, $\beta \in I(\gamma_1,\gamma_2)$ satisfying $\alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4$ we have $\alpha/2$, $\beta/2 \in Z(\delta)$ and $\alpha^2 \equiv 0 \pmod 4$. This implies $\beta^2\gamma_1/\gamma_2 \equiv 0 \pmod 4$ or equivalently $(\beta/2)^2\gamma_1/\gamma_2 \in Z(\delta)$. It follows that $\beta/2 \in I(\gamma_1,\gamma_2)$. Therefore if any of the congruences in (6) was solvable then one of $g/2$, $\sigma/2$, $(g+\sigma)/2$ would be in $I(\gamma_1,\gamma_2)$. This is clearly false and so (5) implies (6). Furthermore it is easily seen that the form of $Z(\delta)(\sqrt{\gamma})$ given in theorem 3.2 reduces to the form given in this theorem when $\alpha/2 \in Z(\delta)$, $\beta/2 \in I(\gamma_1,\gamma_2)$.

Now suppose (6) is true. Given any $\alpha \in Z(\delta)$, $\beta \in I(\gamma_1,\gamma_2)$ such that $\alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4$ we must show that $\alpha \equiv \beta \equiv 0 \pmod 2$.
We have

$$\alpha = 2\kappa + \theta,\ \kappa \in Z(\delta),\ \theta \in \{0,1,\omega,1+\omega\}$$

$$\beta = 2\lambda + \eta,\ \lambda \in I(\gamma_1,\gamma_2),\ \eta \in \{0,g,\sigma,(g+\sigma)/2\}$$

It is easily checked that $\alpha^2 \equiv \theta^2 \pmod 4$. We also have

$$\beta^2\gamma_1/\gamma_2 = 4(\lambda^2\gamma_1/\gamma_2 + \lambda\eta\gamma_1/\gamma_2) + \eta^2\gamma_1/\gamma_2$$

Now $\gamma_2 | \lambda^2\gamma_1, \gamma_2 | \eta^2\gamma_1$ implies $\gamma_2^2 | (\lambda\eta\gamma_1)^2$ which implies $\gamma_2 | \lambda\eta\gamma_1$. Thus $\beta^2\gamma_1/\gamma_2 \equiv \eta^2\gamma_1/\gamma_2 \pmod 4$ and it now follows that $\theta^2 \equiv \eta^2\gamma_1/\gamma_2 \pmod 4$. Since none of the congruences in (6) is solvable we have $\eta = 0$ which

implies $\theta^2 \equiv 0 \pmod 4$ and so $\theta = 0$. This of course implies $\alpha \equiv \beta \equiv 0 \pmod 2$. //

The above theorem gives a relatively simple test for determining precisely when (5) holds. However the following partial result is also most useful.

THEOREM 3.7

(a) Let $\gamma \in Z(\delta)$. Then

$$\forall \alpha,\beta \in Z(\delta),\ \alpha^2 \equiv \beta^2\gamma \pmod 4 \text{ implies } \alpha \equiv \beta \equiv 0 \pmod 2$$

if and only if

(i) $d \equiv 1 \pmod{16}$ and $\gamma \equiv 2,3,(5+\delta)/2,$ or $(1+3\delta)/2 \pmod 4$

or (ii) $d \equiv 5 \pmod{16}$ and $\gamma \equiv 2,3,(1+\delta)/2,(5+\delta)/2,(7+\delta)/2,\delta,1+\delta,2+\delta,3+\delta,(1+3\delta)/2,(3+3\delta)/2,$ or $(5+3\delta)/2 \pmod 4$

or (iii) $d \equiv 9 \pmod{16}$ and $\gamma \equiv 2,3,(1+\delta)/2,$ or $(5+3\delta)/2 \pmod 4$

or (iv) $d \equiv 13 \pmod{16}$ and $\gamma \equiv 2,3,(1+\delta)/2,(3+\delta)/2,(5+\delta)/2,\delta,1+\delta,2+\delta,3+\delta,(1+3\delta)/2,(5+3\delta)/2,$ or $(7+3\delta)/2 \pmod 4$

or (v) $d \equiv 2,3 \pmod 4$ and $\gamma \equiv \delta,1+\delta,2+\delta,3+\delta,3\delta,1+3\delta,2+3\delta,$ or $3+3\delta \pmod 4$

(b) Let γ,γ_1,γ_2 be as in theorem 3.2 and suppose d,γ satisfy one of the congruences in part (a). Then

$$Z(\delta)(\sqrt{\gamma}) = \{\alpha+(\beta/\gamma_2)\sqrt{\gamma} : \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2)\}$$

PROOF

(a) The proof is straightforward but tedious. There are 96 cases to consider. (16 residues for γ modulo 4 for each of the 6 classes of d). Example 3.1 illustrates the procedure for testing each case.

(b) Suppose γ satisfies one of the congruences in part (a). Let $\alpha \in Z(\delta)$, $\beta \in I(\gamma_1,\gamma_2)$ satisfy $\alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4$. Then $(\alpha\gamma_2)^2 \equiv \beta^2\gamma \pmod 4$ and $\alpha^2\gamma \equiv (\beta\gamma_1)^2 \pmod 4$. Therefore by part (a) we have $\alpha\gamma_2 \equiv \beta \equiv 0 \pmod 2$ and $\alpha \equiv \beta\gamma_1 \equiv 0 \pmod 2$. Thus $\alpha \equiv \beta \equiv 0 \pmod 2$ and so (5) holds. The result now follows from theorem 3.6. //

The congruences in theorem 3.7 give sufficient but not necessary conditions for (5) to be true, that is it is possible to have (5) true but γ not satisfying one of the congruences of theorem 3.7. (For example $d = -5$, $\gamma = \gamma_2 = 2, \gamma_1 = 1$). Hence the description of theorem 3.7 as a partial result.

We have now developed a number of conditions which enable us to determine when the form of $Z(\delta)(\sqrt{\gamma})$ given in theorem 3.2 simplifies. The greatest simplification occurs when two of these conditions hold simultaneously.

THEOREM 3.8

Let γ be a non-square rational square-free $Z(\delta)$ integer. Then

$$Z(\delta)(\sqrt{\gamma}) = \{\alpha + \beta\sqrt{\gamma} : \alpha,\beta \in Z(\delta)\}$$

if and only if

$$\begin{aligned} &\gamma \text{ satisfies one of the congruences in theorem 3.7} \\ &\text{and } N(\gamma) \text{ satisfies condition (c) of theorem 3.3} \end{aligned} \tag{7}$$

PROOF

If γ satisfies the conditions in (7) then theorem 3.4, 3.7 show that $Z(\delta)(\sqrt{\gamma})$ is of the stated form.

Now suppose γ does not satisfy the conditions in (7). If γ does not satisfy one of the congruences in theorem 3.7 then $\exists\, \kappa,\lambda \in Z(\delta)$ not

both 0 (mod 2) such that $\kappa^2 \equiv \lambda^2\gamma \pmod 4$. Clearly $(\kappa+\lambda\sqrt{\gamma})/2 \in Z(\delta)(\sqrt{\gamma})$ is not of the form $\alpha+\beta\sqrt{\gamma}$. If $N(\gamma)$ does not satisfy condition (c) in theorem 3.3 then the proof of that theorem shows that $\exists\ \lambda \in Z(\delta)$ such that $\gamma \mid \lambda^2, \gamma \nmid \lambda$. Thus $(\lambda/\gamma)\sqrt{\gamma} \in Z(\delta)(\sqrt{\gamma})$ is not of the form $\alpha+\beta\sqrt{\gamma}$. //

DEFINITION 3.2

Let γ be a non-square rational square-free $Z(\delta)$ integer. Then $Z(\delta)(\sqrt{\gamma})$ can be described as being of one or more of the following forms.

(a) If we can express $Z(\delta)(\sqrt{\gamma})$ as

$$Z(\delta)(\sqrt{\gamma}) = \{\alpha+\beta\sqrt{\gamma} : \alpha,\beta \in Z(\delta)\}$$

that is γ satisfies condition (7) in theorem 3.8 then we say that $Z(\delta)(\sqrt{\gamma})$ is of *form 1*.

(b) If we can express $Z(\delta)(\sqrt{\gamma})$ as

$$Z(\delta)(\sqrt{\gamma}) = \{(\alpha+\beta\sqrt{\gamma})/2 : \alpha,\beta \in Z(\delta),\ \alpha^2 \equiv \beta^2\gamma \pmod 4\}$$

that is $\gamma_1 = \gamma$ satisfies condition (c) of theorem 3.3 then we say that $Z(\delta)(\sqrt{\gamma})$ is of *form 2*.

(c) If we can express $Z(\delta)(\sqrt{\gamma})$ as

$$Z(\delta)(\sqrt{\gamma}) = \{\alpha+(\beta/\gamma_2)\sqrt{\gamma} : \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2)\}$$

that is $\gamma, \gamma_1, \gamma_2$ satisfy condition (5) of theorem 3.6 then we say that $Z(\delta)(\sqrt{\gamma})$ is of *form 3*.

(d) If we can express $Z(\delta)(\sqrt{\gamma})$ as

$$Z(\delta)(\sqrt{\gamma}) = \{(\alpha+(\beta/\gamma_2)\sqrt{\gamma})/2 : \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2),\ \alpha^2 \equiv \beta^2\gamma_1/\gamma_2 \pmod 4\}$$

then we say that $Z(\delta)(\sqrt{\gamma})$ is of *form 4*. //

Of course all $Z(\delta)(\sqrt{\Upsilon})$ are of form 4 (see theorem 3.2) and many are of one or more of the simpler forms. For example if Υ satisfies (7) then $Z(\delta)(\sqrt{\Upsilon})$ is of forms 1, 2, 3, 4. Note that the form(s) to which $Z(\delta)(\sqrt{\Upsilon})$ belongs depends on Υ and it is possible to have $Z(\delta)(\sqrt{\Upsilon}) = Z(\delta)(\sqrt{\kappa})$ with $Z(\delta)(\sqrt{\Upsilon})$, $Z(\delta)(\sqrt{\kappa})$ belonging to distinct sets of forms. Of course we will normally think of $Z(\delta)(\sqrt{\Upsilon})$ as belonging principally to the simplest of the forms to which it belongs.

The above results give us sufficient information about $Z(\delta)(\sqrt{\Upsilon})$ for our purposes. Of course these results do not give us a particularly precise description of those cases where (5) does not hold. However if a more precise description of such a case is required then it is not particularly difficult to carry out a detailed investigation of congruence (2).

The remainder of this section is devoted to the development of a classification of quartic fields $Q(\Gamma)$ of the form $Q(\delta)(\sqrt{\Upsilon})$ and to the presentation of some of the basic results concerning the unit group structure of the various resulting types of field. Note that a more detailed study of this topic can be found in Nagell [1961]. However the simplified version which follows is quite sufficient for our purposes.

Quartic fields which have a quadratic subfield can be divided into two major types.

Type I - those quartic fields which have a complex quadratic subfield.

Type II - those quartic fields which have a real quadratic subfield but no complex quadratic subfield.

(Our type I quartic fields correspond to Nagell's classes 4, 10 fields and our type II quartic fields correspond to Nagell's classes 5, 6, 7, 8, 9, 11 fields). Later in this section we shall subdivide the type II

quartic fields into 3 subtypes. However we first look at type I quartic fields in order to determine their unit group structure.

TYPE I QUARTIC FIELDS

These fields are of the form $Q(\delta)(\sqrt{\gamma})$ where $Q(\delta)$ is a complex quadratic field and γ is a non-square rational square-free $Z(\delta)$ integer.

THEOREM 3.9

The units of the type I quartic field $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ are of the form

$$\xi^j U_f^k, \quad 1 \leq j \leq u, \; k \in Z$$

where U_f is a fundamental unit of $Q(\Gamma)$, that is a unit U for which $|\ell n|U||$ is non-zero and minimal, and ξ is a primitive u^{th} root of unity as described below.

(a) If $Q(\Gamma) \not\supseteq Q(i)$, $Q(\sqrt{-3})$ then $u = 2$ and we have $\xi = -1$.

(b) If $Q(\Gamma) \supseteq Q(i)$, $Q(\Gamma) \not\supseteq Q(\sqrt{-3})$, $Q(\Gamma) \neq Q(\sqrt{i})$ then $u = 4$ and we can take $\xi = i$.

(c) If $Q(\Gamma) \supseteq Q(\sqrt{-3})$, $Q(\Gamma) \not\supseteq Q(i)$, $Q(\Gamma) \neq Q(\sqrt{(1+\sqrt{-3})/2})$ then $u = 6$ and we can take $\xi = (1+\sqrt{-3})/2$.

(d) If $Q(\Gamma) = Q(\sqrt{i})$ then $u = 8$ and we can take $\xi = \sqrt{i}$.

(e) If $Q(\Gamma) = Q(i)(\sqrt{-3})$ then $u = 12$ and we can take $\xi = \sqrt{(1+\sqrt{-3})/2}$.

PROOF

If $\gamma \notin Z$ then from theorem 3.1 we have $Q(\Gamma) = Q(\sqrt{\gamma})$. Now $\sqrt{\gamma}$ has conjugates $-\sqrt{\gamma}$, $\pm\sqrt{\gamma'}$ and each of these four numbers is non-real since $\gamma \notin \mathbb{R}$. If however $\gamma \in Z$ then from theorem 3.1 we have $Q(\Gamma) = Q(\delta+\sqrt{\gamma})$. The conjugates of $\delta + \sqrt{\gamma}$ are $\delta-\sqrt{\gamma}$, $-\delta \pm \sqrt{\gamma}$. These four conjugates are non-real since δ is pure imaginary and $\delta = \pm\sqrt{\gamma}$ is not allowed. Thus in all cases $Q(\Gamma)$ and its conjugate fields are non-real. Dirichlet's theorem (theorem 1.1) and the notion of the regulator now give the theorem apart from the specific form of the roots of unity. Results on

the roots of unity of $Q(\Gamma)$ can be found in Nagell [1961, p356]. //

In section two of this chapter we shall show how a fundamental unit of a type I quartic field can be calculated.

TYPE II QUARTIC FIELDS

These fields are of the form $Q(\delta)(\sqrt{\gamma})$ where $Q(\delta)$ is a real quadratic field and γ is a non-square rational square-free $Z(\delta)$ integer. Furthermore $Q(\delta)(\sqrt{\gamma})$ does not have a complex quadratic subfield. Therefore $\gamma \neq -k\alpha^2$ where $k \in Z^+$, $\alpha \in Q(\delta)$.

Type II quartic fields can be divided into three subtypes according to the signs of γ and its conjugate γ'.

Type IIa $\quad \gamma, \gamma' < 0$

Type IIb $\quad \gamma\gamma' < 0$

Type IIc $\quad \gamma, \gamma' > 0$

We now look briefly at each of these three subtypes.

THEOREM 3.10

The units of the type IIa quartic field $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ are of the form

$$\xi^j \varepsilon(d)^k, \ 1 \leq j \leq u, \ k \in Z$$

where $\varepsilon(d)$ is the fundamental unit of the real quadratic subfield $Q(\delta)$, and ξ is a primitive u^{th} root of unity as follows. If $Q(\Gamma) = Q(\sqrt{-5})(\sqrt{(-5+\sqrt{5})/2})$ then $u = 10$ and we can take $\xi = ((1+\sqrt{5})/2 + \sqrt{(-5+\sqrt{5})/2})/2$. In all other cases we have $u = 2$, $\xi = -1$.

PROOF

Since $\gamma < 0$ and $\gamma \neq -k\alpha^2 (k \in Z^+, \alpha \in Q(\delta))$ it follows that $\gamma \notin Z$. Therefore $Q(\Gamma) = Q(\sqrt{\gamma})$. We also have $\gamma' < 0$. Consequently $\sqrt{\gamma}$ and

its three conjugates $-\sqrt{\Upsilon}$, $\pm\sqrt{\Upsilon'}$ are all non-real. It therefore follows from Dirichlet's theorem that the units of $Q(\Gamma)$ are of the form $\xi^j U^k$ where ξ generates the roots of unity in $Q(\Gamma)$ and U is a fundamental unit of $Q(\Gamma)$. The specific details about ξ and U can be found in Nagell [1961, pp356-359]. //

Thus with one exception we see that the units of a type IIa quartic field are precisely those of the real quadratic subfield. In all cases a fundamental unit can be calculated using the simple continued fraction algorithm and so no further comment is necessary.

THEOREM 3.11

The units of the type IIb quartic field $Q(\Gamma) = Q(\delta)(\sqrt{\Upsilon})$ are of the form

$$\pm U_1^j U_2^k, \quad j,k \in Z$$

where $\{U_1, U_2\}$ is a fundamental system of units of $Q(\Gamma)$, that is a pair of units for which

$$\begin{vmatrix} \ell n|U_1| & \ell n|\sigma(U_1)| \\ \ell n|U_2| & \ell n|\sigma(U_2)| \end{vmatrix}$$

is non-zero and of minimal magnitude where σ is a monomorphism from $Q(\Gamma)$ to $\mathbb{C}$ with $\sigma(U_1) \neq U_1, \bar{U}_1$.

PROOF

We have $\Upsilon\Upsilon' < 0$. Thus Υ, Υ' are of opposite sign and so $\Upsilon \notin Z$. We therefore have $Q(\Gamma) = Q(\sqrt{\Upsilon})$. The fact that Υ, Υ' are of opposite sign also implies that two of the four conjugates $\pm\sqrt{\Upsilon}, \pm\sqrt{\Upsilon'}$ are real while the other two are pure imaginary. The theorem now follows from Dirichlet's theorem and the notion of the regulator. //

In section three of this chapter we shall consider the problem of calculating a fundamental system of units for a type IIb quartic field. Note that in practice we can restrict our attention to real type IIb quartic fields. This is because $Q(\delta)(\sqrt{\gamma})$, $Q(\delta)(\sqrt{\gamma'})$ are isomorphic fields of which one is real and the other is complex when $\gamma\gamma' < 0$.

THEOREM 3.12

The units of the type IIc quartic field $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ are of the form

$$\pm U_1^j U_2^k U_3^\ell, \quad j,k,\ell \in Z$$

where $\{U_1,U_2,U_3\}$ is a fundamental system of units of $Q(\Gamma)$, that is a triple of units for which

$$\begin{vmatrix} \ell n|U_1| & \ell n|\sigma_1(U_1)| & \ell n|\sigma_2(U_1)| \\ \ell n|U_2| & \ell n|\sigma_1(U_2)| & \ell n|\sigma_2(U_2)| \\ \ell n|U_3| & \ell n|\sigma_1(U_3)| & \ell n|\sigma_2(U_3)| \end{vmatrix}$$

is non-zero and of minimal magnitude where σ_1,σ_2 are any two of the three non-identity monomorphisms from $Q(\Gamma)$ to $\mathbb{C}$.

PROOF

Since $\gamma,\gamma' > 0$ we have $\sqrt{\gamma}$, $\sqrt{\gamma'} \in \mathbb{R}$. Therefore $Q(\Gamma)$ and its conjugate fields are all real fields. The result now follows from Dirichlet's theorem and the notion of the regulator. //

We shall not be presenting any further results concerning type IIc quartic fields. However further information on this type of field can be found in the literature. For example Kuroda [1943] and Kubota [1956] have considered type IIc quartic fields of the form $Q(\delta)(\sqrt{d_1})$, d_1 a positive square-free integer. Nagell [1961] gives some general results

concerning type IIc quartic fields. (Nagell's classes 8, 9, and 11 fields). More recently Levesque [1981] and Frei [1982] have given fundamental systems of units for certain subclasses of fields of the form $Q(\delta)(\sqrt{d_1})$.

SECTION TWO

UNITS OF TYPE I QUARTIC FIELDS

In this section we consider the problem of calculating fundamental units of type I quartic fields, that is fields which are quadratic extensions of complex quadratic fields. (See section one of this chapter). We have previously indicated that the work in chapter two was largely motivated by the desire to calculate fundamental units of this type of field. However we shall see that algorithm 2.2 or a slightly modified version of algorithm 2.2 are only guaranteed to find a fundamental unit of a type I quartic field $Q(\delta)(\sqrt{\gamma})$ when $Z(\delta)(\sqrt{\gamma})$ is of form 1 or form 2. Thus the work in chapter two has only been partially successful in achieving one of its main objectives. Of course the problem lies in the fact that $Z(\delta)(\sqrt{\gamma})$ and $Z(\delta)[1,\sqrt{\gamma}]$ (which is isomorphic to $M(\sqrt{\gamma})$) are significantly different when $Z(\delta)(\sqrt{\gamma})$ is not of form 1 or form 2. When $Z(\delta)(\sqrt{\gamma})$ is of form 3 or form 4 algorithm 2.2 or its modified version will certainly produce a unit $U \in Z(\delta)(\sqrt{\gamma})$. However U may or may not be fundamental. We will therefore develop a procedure which will either verify that U is fundamental or locate a fundamental unit of the form $(\xi U)^{1/m}$ where ξ is a root of unity and $m \in Z^+$, $m \geq 2$. The ideas used are basically those used in Jeans and Hendy [1978] and Jeans [1978]. This will give a satisfactory method for finding fundamental units when $Z(\delta)(\sqrt{\gamma})$ is of

form 3 or form 4. However we will also briefly note an alternative idea for dealing with these cases. Although this idea has not yet been completely developed it appears likely that this alternative approach will eventually result in a far more satisfactory method for calculating fundamental units for the form 3 and form 4 cases. To illustrate the use of the methods developed in this section we shall calculate fundamental units for the distinct non-isomorphic type I quartic fields $Q(\sqrt[4]{D})$, $D \in Z$, $-99 \leq D \leq -1$. We then finish this section by reviewing some of the relevant literature.

Throughout this section we assume that $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ is a type I quartic field, that is $Q(\delta)$ is a complex quadratic field and γ is a non-square rational square-free $Z(\delta)$ integer. We shall use U_f to denote a fundamental unit of $Q(\Gamma)$ which satisfies $|U_f| > 1$. Note that if U is a unit of $Q(\Gamma)$ then $N_\delta(U)$ is a $Z(\delta)$ root of unity.

There is a trivial isomorphism between

$$S = \{(\alpha + \beta\sqrt{\gamma},\ \alpha - \beta\sqrt{\gamma}) : \alpha,\beta \in Q(\delta)\} \subseteq \mathbb{C} \times \mathbb{C}$$

and $Q(\delta)(\sqrt{\gamma})$. To simplify the presentation of results in this section we therefore drop the two dimensional notation of chapter two. Thus $A = (\alpha + \beta\sqrt{\gamma},\ \alpha - \beta\sqrt{\gamma})$ becomes $A = \alpha + \beta\sqrt{\gamma}$, $M(\sqrt{\gamma})$ and $Z(\delta)[1,\sqrt{\gamma}]$ are now identical as are A^* (A reverse) and A^* (A conjugate), and N_δ now has the standard meaning, that is the relative norm function from $Q(\delta)(\sqrt{\gamma})$ to $Q(\delta)$.

We shall successively develop four algorithms for the calculation of fundamental units of type I quartic fields $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$. Algorithm 3.j, $j = 1,2,3,4$ will apply to the calculation of U_f when $Z(\delta)(\sqrt{\gamma})$ is of form j. Since $Z(\delta)(\sqrt{\gamma})$ can be of more than one form there will often be several algorithms which could be used to calculate

U_f for a given $Q(\Gamma)$. However in practice it will obviously pay from an efficiency point of view to use one of the simpler algorithms (smaller j) whenever possible.

We begin by considering the case where $Z(\delta)(\sqrt{\gamma})$ is of form 1. In chapter two section four we saw that the calculation of a period of relative minima of $M(\sqrt{\gamma})$ produces a unit $A_q \in Z(\delta)(\sqrt{\gamma})$. (Since $|\sqrt{\gamma}| \geq 1$ we have $A_0 = 1$). When $Z(\delta)(\sqrt{\gamma})$ is of form 1 we can guarantee that A_q is fundamental.

THEOREM 3.13

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ be a type I quartic field for which $Z(\delta)(\sqrt{\gamma})$ is of form 1. Then the set of fundamental units of $Q(\Gamma)$ having magnitude greater than 1 is E_q where q is the unique length of a period of relative minima of $M(\sqrt{\gamma})$.

PROOF

If $Z(\delta)(\sqrt{\gamma})$ is of form 1 then $Z(\delta)(\sqrt{\gamma})$ and $M(\sqrt{\gamma})$ are identical. Let U_f, $|U_f| > 1$ be a fundamental unit of $Q(\Gamma)$. Then $U_f \in M(\sqrt{\gamma})$. Now $A \in M(\sqrt{\gamma})$ implies $N_\delta(A) \in Z(\delta)$ and so if $A \neq 0$ then $|N_\delta(A)| \geq 1$. Since $|N_\delta(U_f)| = 1$ it is easily seen that U_f is a relative minimum of $M(\sqrt{\gamma})$. The fact that U_f is fundamental ensures that $M(\sqrt{\gamma})$ has a minimal period of relative minima $A_0, A_1, \ldots, A_q = U_f$. It is easily checked that if $A_q^{(1)} \sim A_q$ then $A_q^{(1)}$ is also a fundamental unit of $Q(\Gamma)$ and so the theorem is clear. //

EXAMPLE 3.2

(a) Let $d = -10$, $\gamma = 1+\delta$. We have previously seen (examples 2.1, 2.3, 2.5) that $M(\sqrt{\gamma})$ has a period of length 6 with

$$A_6 = (245-20\delta) + (88-34\delta)\sqrt{\gamma}$$

It is a simple matter to check that γ satisfies condition (7) in theorem 3.8 and so $Z(\delta)(\sqrt{\gamma})$ is of form 1. Theorem 3.13 therefore allows us to

conclude that A_6 is a fundamental unit of $Q(\delta)(\sqrt{Y})$.

(b) In the same manner as part (a) we reach the conclusion

$$A_5 = (62-25\delta) + ((-25-9\delta)/2)\sqrt{-3+5\delta}$$

is a fundamental unit of $Q(\delta)(\sqrt{-3+5\delta})$, $\delta = \sqrt{-35}$. (See example 2.7b). //

We now define algorithm 3.1 which is basically a simplified version of algorithm 2.2.

ALGORITHM 3.1

Let $Q(\Gamma) = Q(\delta)(\sqrt{Y})$ be a type I quartic field for which $Z(\delta)(\sqrt{Y})$ is of form 1. Then a fundamental unit U_f of $Q(\Gamma)$ can be calculated using algorithm 2.2 modified as follows.

(a) We have $h = 1$ and so all steps involving h can be simplified accordingly. In particular step 2 of algorithm 2.2 can be omitted since $|W| = |\sqrt{Y}| \geq 1$.

(b) Note that as a consequence of the simplified notation now being used we have

$$W = w = \sqrt{Y}, \quad W^* = -\sqrt{Y}$$

$$W_k = w_{k,1} = (\theta_k + g_k\sqrt{Y})/N_\delta(A_k), \quad W_k^* = w_{k,2} = (\theta_k - g_k\sqrt{Y})/N_\delta(A_k)$$

However the simplified notation causes one problem in that step 19 of algorithm 2.2B no longer calculates W_{k+1}^*. We must therefore add the calculation

$$W_{k+1}^* = (\eta_{k,3}+\eta_{k,4}W_k^*)/(\eta_{k,1}+\eta_{k,2}W_k^*)$$

to this step. Step 22 of algorithm 2.2B must also be modified in a similar manner.

(c) Step 17 of algorithm 2.2 can be reduced to

$$U_f = A_q = \begin{cases} A_{k+1}^{(j)}/A_k^* & q \text{ odd} \\ A_{k+1}^{(j)}/(A_{k+1}^{(1)})^* & q \text{ even} \end{cases}$$ //

If $Z(\delta)(\sqrt{\Upsilon})$ is not of form 1 then

$$M(\sqrt{\Upsilon}) = Z(\delta)[1,\sqrt{\Upsilon}] \subsetneq Z(\delta)(\sqrt{\Upsilon})$$

Consequently the unit A_q obtained from a period of $M(\sqrt{\Upsilon})$ may or may not be fundamental. We now consider how algorithm 2.2 can be modified to overcome this problem when $Z(\delta)(\sqrt{\Upsilon})$ is of form 2. Note that the resulting algorithm 3.2 plus algorithm 3.1 will effectively allow us to calculate U_f for any type I quartic field $Q(\delta)(\sqrt{\Upsilon})$ for which $h(d) = 1$. (See the comments following theorem 3.4).

THEOREM 3.14

Let $Q(\Gamma) = Q(\delta)(\sqrt{\Upsilon})$ be a type I quartic field for which $Z(\delta)(\sqrt{\Upsilon})$ is of form 2. Let A_q be a unit of $M(\sqrt{\Upsilon})$ corresponding to a minimal period and suppose A_q is not fundamental. Then we can find $V_f \in M(\sqrt{\Upsilon})$ satisfying

(i) $N_\delta(V_f) = 4\xi$, ξ a $Z(\delta)$ root of unity

(ii) $2 < |V_f| \leq 2\sqrt{|A_q|}$, $|V_f|$ minimal

Any such V_f gives $V_f/2$ as a fundamental unit of $Q(\Gamma)$. Let $m \in Z^+ \cup \{0\}$ satisfy $|A_m| \leq |V_f| < |A_{m+1}|$. Then

either $V_f \sim A_m$, $m \leq [q/2]$

or $V_f \not\sim A_m$ and $|N_\delta(A_m)| < 4$

PROOF

Since A_q is not fundamental we have $A_q = \xi_1 U_f^n$, ξ_1 a $Z(\delta)(\sqrt{\Upsilon})$ root of unity, U_f a fundamental unit with $|U_f| > 1$, and $n \geq 2$. Clearly $V_f = 2U_f$ satisfies (i), (ii).

Given any $V_f = \alpha + \beta\sqrt{\Upsilon} \in M(\sqrt{\Upsilon})$ satisfying (i), (ii) we have $N_\delta(V_f/2) = \xi$, $(V_f/2) + (V_f^*/2) = \alpha$ and so $V_f/2$ is a unit of $Z(\delta)(\sqrt{\Upsilon})$ which in view of the minimality of $|V_f|$ is clearly fundamental.

Suppose $V_f \sim A_m$. Then $V_f = A_m^{(1)}$ with $N_\delta(A_m^{(1)}) = 4\xi$ and clearly $m < q$. Now $A_{q-m} = A_q(A_m^{(1)})^*$ satisfies $N_\delta(A_{q-m}) = 4\xi_2$, ξ_2 a $Z(\delta)$ root of unity, $|A_{q-m}| > |N_\delta(A_{q-m})| > 2$. The minimality of $|V_f|$ now implies $|A_m^{(1)}| \le |A_{q-m}|$, that is $m \le [q/2]$.

Finally suppose $V_f \not\sim A_m$. Since A_m, A_{m+1} are consecutive magnitude minima we have $|V_f^*| \ge |A_m^*|$. Thus

$$|N_\delta(A_m)| = |A_m A_m^*| < |V_f V_f^*| = |N_\delta(V_f)| = 4$$

The inequality is strict since $V_f \not\sim A_m$. //

To find U_f when $Z(\delta)(\sqrt{\Upsilon})$ is of form 2 we basically have to determine whether or not V_f as described in theorem 3.14 exists. (If it does then take $U_f = V_f/2$ otherwise take $U_f = A_q$). If V_f exists (that is A_q is not fundamental) then we generally have $V_f \sim A_m$, $m \le [q/2]$ and so V_f is effectively located by algorithm 2.2. However V_f can be masked from the view of algorithm 2.2 by the presence of a relative minimum of small norm, that is A_m when $|N_\delta(A_m)| < 4$. Consequently calculation of U_f is not quite as simple as we might hope for.

EXAMPLE 3.3

(a) Let $d = -2$, $\gamma = 5+6\delta$. We have $N(\Upsilon) = 97$, $\Upsilon \equiv 1+2\delta \pmod 4$ and it is easily checked that $Z(\delta)(\sqrt{\Upsilon})$ is of form 2 but not of form 1. Algorithm 2.2 produces the following period of length 4.

TABLE 3.1 Period for $M(\sqrt{5+6\delta})$

k	0	1	2	3	4
α_k	1	$2+\delta$	$-3-\delta$	$-3-5\delta$	$6+6\delta$
β_k	0	1	-1	$-2-\delta$	$3+\delta$
$N_\delta(A_k)$	1	$-3-2\delta$	2	$-3-2\delta$	1

We have $A_4 = -U_f^2$ where $U_f = ((-2+3\delta) + \delta\sqrt{\Upsilon})/2$ and $V_f = 2U_f$ satisfies $|A_2| < |V_f| < |A_3|$, that is V_f is not a relative minimum. In fact

$V_f = -\delta A_2$. Note that $|N_\delta(A_2)| = 2 < 4$.

(b) Let $d = -11$, $\gamma = (3+\delta)/2$. It is easily checked that $Z(\delta)(\sqrt{\gamma})$ is of form 2 but not of form 1. Algorithm 2.2 produces the following period of length 3.

TABLE 3.2 Period for $M(\sqrt{(3+\delta)/2})$

k	0	1	2	3
α_k	1	1	$(-3+\delta)/2$	$1+\delta$
β_k	0	1	$(-1+\delta)/2$	$(3+\delta)/2$
$N_\delta(A_k)$	1	$(-1-\delta)/2$	$(1+\delta)/2$	-1

We have $A_3 = -U_f^5$ where $U_f = ((1+\delta)/2 + \sqrt{\gamma})/2$ and $|A_1| < |V_f| < |A_2|$. Thus V_f is not a relative minimum and it is easily checked that $V_f \neq \kappa A_1$ for any $\kappa \in Q(\delta)$. //

Recall the standard quadratic case $d_1 \equiv 1 \pmod 4$. There we find that if $\varepsilon(d_1) \notin Z[1,\sqrt{d_1}]$ then $\varepsilon(d_1)$ is still guaranteed to correspond to a simple continued fraction convergent of $\sqrt{d_1}$ provided $d_1 > 5$. (See theorem 1.9). We can prove a similar result for V_f (that is guarantee $V_f \sim A_m$ when A_q is not fundamental) provided $|\sqrt{\gamma}|$ is sufficiently large and $d \neq -1, -2, -7$.

THEOREM 3.15

Let $Q(\Gamma), A_q, V_f, A_m, \xi$ be the same as in theorem 3.14.

(a) Suppose $V_f \not\sim A_m$. Let $T_f = V_f/A_m \in M_m(\sqrt{\gamma})$. Then

$$T_f = (\mu_1+\mu_2 W_m)/g_m,\ (\mu_1,\mu_2)\ W_m \text{ allowable}$$

$$|\mu_2| \leq (4+|N_\delta(A_m)|)/2|\sqrt{\gamma}| = b(\mu_2)$$

$$\mu_1 = (-x+s\sqrt{(x^2-4y)})/2 \qquad (8)$$

where $s \in \{-1,1\}$, $x = \mu_2(W_m+W_m^*)$, $y = \mu_2^2 W_m W_m^* - 4\xi g_m^2/N_\delta(A_m)$. If $\mu_2 = 0$ then $T_f = \mu_1$ (that is $g_m = 1$) and

$$\left.\begin{aligned} &\text{either } d = -1,\ N_\delta(A_m) = 2\delta^j,\ \mu_1 = (1+\delta)\delta^n,\ j,n \in Z,\ m = q/2 \\ &\text{or } d = -2,\ N_\delta(A_m) = \pm 2,\ \mu_1 = \pm\delta,\ m = q/2 \\ &\text{or } d = -7,\ N_\delta(A_m) = \pm(3\pm\delta)/2,\ \mu_1 = \pm(1\pm\delta)/2 \\ &\qquad\qquad \mu_1^2 = \pm N_\delta(A_m)',\ m \le [q/2] \end{aligned}\right\} \quad (9)$$

(b) If $|\sqrt{Y}| \ge 4$ and $d \neq -1, -2, -7$ then $V_f \sim A_m$.

PROOF

(a) T_f is clearly of the stated $M_m(\sqrt{Y})$ form. To prove the bound for $|\mu_2|$ we begin by noting that

$$T_f - T_f^* = \mu_2(W_m - W_m^*)/g_m$$

From theorem 2.26 we have $W_m - W_m^* = 2g_m\sqrt{Y}/N_\delta(A_m)$ and so

$$\mu_2 = N_\delta(A_m)(T_f - T_f^*)/2\sqrt{Y}$$

Since $|A_m| \le |V_f|$, $|A_m^*| \le |V_f^*|$, $N_\delta(T_f)N_\delta(A_m) = N_\delta(V_f) = 4\xi$ (see theorem 3.14 and proof) it is easily checked that

$$1 \le |T_f|, |T_f^*| \le 4/|N_\delta(A_m)|$$

$$|T_f||T_f^*| = 4/|N_\delta(A_m)|$$

and it is then easily deduced that

$$|T_f - T_f^*| \le |T_f| + |T_f^*| \le 4/|N_\delta(A_m)| + 1$$

This gives the bound for $|\mu_2|$. The expression for μ_1 is obtained from the quadratic in μ_1 which can be obtained from the relationship $N_\delta(T_f) = N_\delta(V_f/A_m)$.

Now suppose $\mu_2 = 0$. Then $V_f = (\mu_1/g_m)A_m$ and so taking norms we have $4\xi = (\mu_1/g_m)^2 N_\delta(A_m)$. This implies

$$\alpha = \sqrt{N_\delta(A_m)\xi^{-1}} = \pm 2g_m/\mu_1 \in Q(\delta)$$

Since $N_\delta(A_m)\xi^{-1} \in Z(\delta)$ it follows that $\alpha \in Z(\delta)$. Now if $|N_\delta(A_m)| = 1$

then $|A_q| \leq |A_m|$, $|\mu_1/g_m| = 2$, and so $|V_f| \geq 2|A_q|$ which contradicts (ii) of theorem 3.14. We therefore have $1 < |N_\delta(A_m)| < 4$. This implies $1 < |\alpha| < 2$. It is easily checked that such an α can only exist if $d = -1, -2, -3, -7,$ or -11. Since $h(d) = 1$ in each of these cases it follows that $g_m = 1$. We now have the conditions $d = -1, -2, -3, -7,$ or -11 and

$$\alpha, \mu_1 \in Z(\delta),\ 1 < |\alpha|, |\mu_1| < 2,\ \alpha\mu_1 = \pm 2,\ \mu_1^2 N_\delta(A_m) = 4\xi$$

Apart from the restrictions on m the results in (9) now follow easily. Clearly $m < q$. If $m > [q/2]$ then

$$V = V_f^* A_q = (\mu_1 A_m)^* A_q = \mu_1 A_m^* A_q = \mu_1 A_{q-m}$$

is easily seen to satisfy $N_\delta(V) = 4\xi_1$, ξ_1 a $Z(\delta)$ root of unity, and $2 < |V| < |V_f|$ which contradicts the minimality of $|V_f|$. Finally if $d = -1, -2$ then

$$A_m/A_m^* = A_m^2/N_\delta(A_m) = (\alpha_m^2 + \beta_m^2\gamma + 2\alpha_m\beta_m\sqrt{\gamma})/N_\delta(A_m)$$

$$= (N_\delta(A_m) + 2\beta_m^2\gamma + 2\alpha_m\beta_m\sqrt{\gamma})/N_\delta(A_m)$$

is an element of $M(\sqrt{\gamma})$ since $N_\delta(A_m) | 2$. It now follows easily that A_m/A_m^* is a unit of $M(\sqrt{\gamma})$ and so $m = q/2$.

(b) Suppose $V_f \neq A_m$. Then part (a) gives $|\mu_2| < 1$ since $|\sqrt{\gamma}| \geq 4$ and $|N_\delta(A_m)| < 4$. Thus $\mu_2 = 0$. However part (a) now implies $d = -1, -2,$ or -7 which is a contradiction. //

We now describe how to modify algorithm 2.2 (giving algorithm 3.2) so that U_f can be calculated whenever $Z(\delta)(\sqrt{\gamma})$ is of form 2. We will of course need appropriately modified versions of algorithms 2.2A, 2.2B.

ALGORITHMS 3.2A, 3.2B

These algorithms are obtained from algorithms 2.2A, 2.2B by making the modifications noted in (a),(b) of algorithm 3.1. //

The major modification required in algorithm 2.2 is the addition of appropriate procedures which are guaranteed to locate V_f whenever A_q is not fundamental. As algorithm 3.2 (that is the modified version of algorithm 2.2) calculates a period of relative minima of $M(\sqrt{Y})$ it will need to test for three possibilities

(i) $V_f \sim A_m$

(ii) $V_f \nsim A_m$, $V_f = \mu_1 A_m$ (see (9))

(iii) $V_f \nsim A_m$, $V_f = T_f A_m$, $T_f = (\mu_1 + \mu_2 W_m)/g_m \in M_m(\sqrt{Y})$

Note that possibilities (i), (ii) cannot occur with $m = 0$. In the interests of efficiency we should carry out the appropriate testing at the earliest practical point in the algorithm. We therefore test for possibility (iii) (with $m = k$) prior to the calculation of the $R_k^{(j)} \in M_k(\sqrt{Y})$, and we test for possibility (i) and then possibility (ii) if $d = -7$ (both with $m = k + 1$) immediately following the calculation of the $R_k^{(j)} \in M_k(\sqrt{Y})$ (testing each $A_{k+1}^{(j)} = R_k^{(j)} A_k$). Finally if we reach the midpoint of an even length period when $d = -1, -2$ then we test for possibility (ii) with $m = q/2$.

To test for possibility (i) we first check to see if $|N_\delta(A_{k+1}^{(1)})| = 4$. In general this simple test is sufficient to reject possibility (i). However if the possibility is not rejected then we test to see if any $A_{k+1}^{(j)}$ has norm 4ξ. (Note that we can have $\alpha \in Z(\delta)$, $|\alpha| = 4$, $\alpha \neq 4\xi$). If we find $N_\delta(A_{k+1}^{(j)}) = 4\xi$ then we can take $V_f = A_{k+1}^{(j)}$ since the algorithm will have already checked for $|V_f| < |A_{k+1}^{(j)}|$. From theorem 3.14 we see that testing for possibility (i) can stop once the midpoint of a period is recognised since at that point we have $k + 1 \geq q/2$.

To test for possibility (ii) when $d = -7$ we check for the occurence of $N_\delta(A_{k+1}^{(j)}) = \pm(3 \pm \delta)/2$. (Note that this requires $|N_\delta(A_{k+1}^{(1)})| = 2$). If such a norm occurs then $V = \mu_1 A_{k+1}^{(j)}$ where $\mu_1 = (1 \pm \delta)/2$, $\mu_1^2 = \pm N_\delta(A_{k+1}^{(j)})'$ is

easily seen to satisfy $N_\delta(V) = 4\xi$, $|V| > 2$. Now the algorithm will have already tested for $|V_f| < |A_{k+1}^{(j)}|$ and so if $|V_f| < |V|$ we must have

$$|A_{k+1}^{(j)}| \leq |V_f| < |U_f||V_f| \leq |V| = |\mu_1||A_{k+1}^{(j)}|$$

Since $|\mu_1| = \sqrt{2}$ it follows that $|U_f| \leq \sqrt{2}$. However in such a case we have $2U_f = \alpha+\beta\sqrt{\gamma}$, $|\beta|, |\sqrt{\gamma}| < 4$ and it is easily checked that $2U_f$ would have been previously located by the testing for possibility (iii) when $k = 0$. (See below). Thus if such a V is found then we can take $V_f = V$. From theorem 3.15 we see that testing for this possibility can stop once the midpoint of a period is recognised.

To test for possibility (ii) at the midpoint of an even length period when $d = -1, -2$ we simply check to see if $|N_\delta(A_{q/2}^{(1)})| = 2$. If this is the case then in a similar manner to the previous case we see that we can take $V_f = (1+\delta)A_{q/2}^{(1)}$ if $d = -1$, and $V_f = \delta A_{q/2}^{(1)}$ if $d = -2$.

Testing for possibility (iii) is required when $|N_\delta(A_k)| < 4$ and

$$b(\mu_2) = (4+|N_\delta(A_k)|)/2|\sqrt{\gamma}| \geq 1$$

When this is the case we must test each $\beta \in Z(\delta)^+ \cap I_k$, $|\beta| \leq b(\mu_2)$ to see if α given by (8) ($\mu_2 = \beta, \xi$ any $Z(\delta)$ root of unity, $s = \pm 1$) is a $Z(\delta)$ integer for which $T = (\alpha+\beta W_k)/g_k \in M_k(\sqrt{\gamma}), |TA_k| > 2$, $1 \leq |T| \leq 4/|N_\delta(A_k)|$. (See theorems 3.14, 3.15). Note that there are 12, 8, or 4 tests required for each β depending on whether or not $d = -3, -1$, or $d \neq -3, -1$. While this may appear rather time consuming note that $b(\mu_2) \geq 1$ can only occur when $|N_\delta(A_k)|, |\sqrt{\gamma}| < 4$ and that $b(\mu_2) < 4/|\sqrt{\gamma}|$. Consequently this testing is generally not required and when it is required the number of β which must be considered is small. In particular the 12 and 8 test cases can only occur for a handful of values of γ. (That is $|\sqrt{\gamma}| < 4$ and $d = -3, -1$). We can also show that it suffices to test for possibility (iii) for $k \leq [q/2]$. To see this suppose $V_f = T_f A_m, m > [q/2]$.

If $m \geq q$ then it is easily checked that $|U_f| \leq 2$ and that $2U_f$ will be located when testing for possibility (iii) with $k = 0$. Therefore assume $m < q$. Given any A_{q-m} we have $V_f = TA_{q-m}$, $T \in M_{q-m}(\sqrt{Y})$ and $|N_\delta(A_{q-m})| = |N_\delta(A_m)|$. Now

$$V = V_f^* A_q = T_f^*(A_m^* A_q)$$

is easily seen to satisfy $N_\delta(V) = 4\xi_1$, ξ_1 a $Z(\delta)$ root of unity, $|V| > 2$ and so the minimality of $|V_f|$ requires $|V_f| \leq |V|$. Since $A_{q-m} \sim A_m^* A_q$ it is now easily deduced that

$$1 \leq |T|, |T^*| \leq 4/|N_\delta(A_{q-m})|$$

If we now check with the proof of theorem 3.15a it is clear that such a T will be located in the testing for possibility (iii) when $k = q-m < [q/2]$.

Now suppose the above testing for possibility (iii) produces T satisfying the required conditions. Unfortunately we cannot conclude that the unit $U = TA_k/2$ is fundamental. If $|U_f|$ is small then it is possible that we may have found a unit for which $|U| \geq |U_f|^2$. Since the algorithm will have previously tested for $|V_f| < |A_k|$ (if $k > 0$) this can only occur if

$$|A_k| \leq |V_f| < |U_f||V_f| = 2|U_f|^2 \leq |TA_k|$$

that is $|U_f| \leq |T|$, $|A_k| \leq |V_f| = 2|U_f| \leq 2|T|$. Thus if $|T| < |A_k|/2$ then U is indeed fundamental. However if $|T| \geq |A_k|/2$ (for example when $k = 0$) then we must continue the possibility (iii) testing to see if a smaller magnitude T can be found. Of course if such a T is found then it is retained in place of the original T value and at the completion of this testing we will indeed have $U = TA_k/2$ is fundamental. (If $k > 0$ then the algorithm will have previously tested for the possibilities $V_f \sim A_k$, $V_f = \mu_1 A_k$). (Note that although U is fundamental we do not necessarily have $T = T_f$. Instead we may have

either $T = 2R_{k,q}$, $k \le q/2$, $U = U_f = A_q$ (T_f, V_f not defined)

or $T = R_{k,m}$, $V_f = R_{k,m}A_k = A_m$, $k < m$ (T_f not defined)

or $T = T_f R_{k,m}$, $V_f = T_f R_{k,m} A_k = T_f A_m$, $k < m$

The first case can only occur when $|U_f| = |A_q|$ is small, generally $k = 0$, $q = 1$. In the second and third cases V_f is simply found earlier than we might expect, that is in $M_k(\sqrt{Y})$ with $k < m$).

The required testing for possibility (iii) is now stated more precisely in the form of an algorithm.

ALGORITHM 3.2C

Given a standard representation of $M_k(\sqrt{Y})$ with $|N_\delta(A_k)| < 4$ and $b(\mu_2) \ge 1$ (plus the fact that if V_f exists then $|V_f| \ge |A_k|$, $V_f \nmid A_k$, $V_f \ne \mu_1 A_k$) then we

either locate $T \in M_k(\sqrt{Y})$ for which $U = TA_k/2$ is fundamental

or verify that if V_f exists then

$$|V_f| \ge \max\{|A_{k+1}|, (4/|N_\delta(A_k)|)|A_k|\}$$

as follows.

1 Set $mT = 4/|N_\delta(A_k)|$, $\beta = g_k$, $m_3 = 1$, $m_4 = 0$

2 If $d = -3$ set $\xi_d = (1+\sqrt{-3})/2$ and go to 5

3 If $d = -1$ set $\xi_d = i$ and go to 5

4 Set $\xi_d = -1$

5 If $|\beta| > b(\mu_2)$ go to 16

6 Set $\xi = \xi_d$, $s = 1$

7 Calculate $x = \beta(W_k + W_k^*)$, $y = \beta^2 W_k W_k^* - 4\xi g_k^2/N_\delta(A_k)$, $z = \sqrt{x^2-4y}$

8 Calculate $v = (-x+sz)/2 - m_4\psi_k$, $m_2 = \{-c\,\mathrm{Im}(v)/|\delta|\}$

$$m_1 = \{\mathrm{Re}(v-m_2\sigma_k')/g_k\}, \quad \alpha = m_1 g_k + m_2\sigma_k' + m_4\psi_k$$

9 If $N_\delta(A_k)N_\delta((\alpha+\beta W_k)/g_k) \neq 4\xi$ go to 14

10 Calculate $t = |(\alpha+\beta W_k)/g_k|$

11 If $t|A_k| \leq 2$ or $t < 1$ or $t > mT$ go to 14

12 Set $\mu_1 = \alpha$, $\mu_2 = \beta$, $T = (\mu_1+\mu_2 W_k)/g_k$, $mT = t$

13 If $mT < |A_k|/2$ then stop

14 If $s = 1$ set $s = -1$ and go to 8

15 If $\xi \neq 1$ multiply ξ by ξ_d, set $s = 1$, and go to 7

16 Increment m_3 by 1

17 Set $\beta = m_3 g_k + m_4 \sigma_k$

18 If $|\beta| \leq b(\mu_2)$ go to 6

19 If $m_4 > 0$ reset m_4 to $-m_4$ and go to 22

20 Reset m_4 to $-m_4 + 1$

21 If $m_4|\delta|/c > b(\mu_2)$ then stop

22 Set $m_3 = -[m_4 a_k/cg_k]$

23 If $m_4 < 0$ and $m_3 g_k + m_4 a_k/c = 0$ then increment m_3 by 1

24 Go to 17 //

Algorithm 3.2 will initialize $T = 0$ and so will recognize the fact that T has been found by testing for $T \neq 0$. (See steps 2.5, 26 of algorithm 3.2). ξ_d in step 2, 3, or 4 is an appropriate primitive $Z(\delta)$ root of unity and steps 14, 15 control the looping through the required 12, 8, or 4 tests for each β. Steps 7, 8, 9 give an effective method for testing whether or not $\alpha = (-x+sz)/2$ is a $Z(\delta)$ integer for which (α,β) is W_k allowable. (This method partially reflects the fact that calculations performed by the algorithm will involve finite precision approximations to irrational numbers).

Of course the final point to note concerning the testing described over the past few pages is that if all testing for possibilities (i), (ii), (iii) fails then we can conclude that A_q is fundamental.

The ideas presented concerning the form 2 cases are now tied together by the following algorithm.

ALGORITHM 3.2

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ be a type I quartic field for which $Z(\delta)(\sqrt{\gamma})$ is of form 2. Then a fundamental unit U_f of $Q(\Gamma)$ can be calculated as follows.

1 Set $A_0 = 1$, $N_\delta(A_0) = 1$, $W_0 = \sqrt{\gamma}$, $W_0^* = -\sqrt{\gamma}$, $g_0 = 1$, $\psi_0 = 0$, $\theta_0 = 0$, $\sigma_0 = \omega$, $T = 0$, $k = 0$

2.1 If $|N_\delta(A_k)| \geq 4$ or $|\sqrt{\gamma}| \geq 4$ go to 3

2.2 Calculate $b(\mu_2) = (4+|N_\delta(A_k)|)/2|\sqrt{\gamma}|$

2.3 If $b(\mu_2) < 1$ go to 3

2.4 Apply algorithm 3.2C

2.5 If $T \neq 0$ set $U_f = TA_k/2$ and stop

3 Use algorithm 3.2A to calculate the $R_k^{(j)} \in M_k(\sqrt{\gamma})$

4.1 If $|N_\delta(A_{k+1}^{(1)})| \neq 4$ go to 5.1

4.2 If $N_\delta(A_{k+1}^{(j)}) = 4\xi$ set $U_f = A_{k+1}^{(j)}/2$ and stop

5.1 If $d \neq -7$ or $|N_\delta(A_{k+1}^{(1)})| \neq 2$ go to 6

5.2 If $N_\delta(A_{k+1}^{(j)}) = \pm(3\pm\delta)/2$ set $U_f = \mu_1 A_{k+1}^{(j)}/2$ where $\mu_1 = (1\pm\delta)/2$ is chosen so that $\mu_1^2 = \pm N_\delta(A_{k+1}^{(j)})'$ and stop

6-16 are the same as in algorithm 2.2 apart from changing "2.2B" to "3.2B" in step 9, and "go to 5" to "go to 2.1" in step 16

17 If q is odd set $U_f = A_q = A_{k+1}^{(j)}/A_k^*$ and stop

18 If $d \neq -1, -2$ or $|N_\delta(A_{k+1}^{(1)})| \neq 2$ go to 20

19 If $d = -1$ then set $U_f = (1+\delta)A_{k+1}^{(1)}/2$ and stop

else set $U_f = \delta A_{k+1}^{(1)}/2$ and stop

20 If $|N_\delta(A_{k+1}^{(1)})| \geq 4$ or $|\sqrt{\Upsilon}| \geq 4$ set $U_f = A_q = A_{k+1}^{(j)}/(A_{k+1}^{(1)})^*$ and stop

21 Set $A_{k+1} = A_{k+1}^{(1)}$, $M_{k+1}(\sqrt{\Upsilon}) = M_{k+1}^{(1)}(\sqrt{\Upsilon})$, $A = A_{k+1}^{(j)}$, $B = A_{k+1}^{(1)}$

22 Increment k by 1

23 Calculate $b(\mu_2) = (4+|N_\delta(A_k)|)/2|\sqrt{\Upsilon}|$

24 If $b(\mu_2) < 1$ set $U_f = A/B^*$ and stop

25 Apply algorithm 3.2C

26 If $T \neq 0$ then set $U_f = TA_k/2$ and stop

else set $U_f = A/B^*$ and stop //

Steps 6-16 are the period midpoint testing steps of algorithm 2.2. If we reach the midpoint of an odd length period then the current value of k satisfies $k+1 = (q+1)/2 > [q/2]$. Consequently all necessary testing for possibilities (i), (ii), (iii) has been completed without locating V_f. We can therefore take $U_f = A_q$. (See step 17). However if we reach the midpoint of an even length period then we must test for possibility (ii) if $d = -1, -2$. Furthermore at the midpoint of an even length period we have $k+1 = q/2$ and so we may still have to test for possibility (iii) with $k = q/2$. Steps 18-26 carry out the necessary testing for these remaining possibilities and set U_f accordingly.

We now briefly illustrate the use of algorithm 3.2.

EXAMPLE 3.4

(a) Let $d = -2$, $\Upsilon = 5+6\delta$. (See example 3.3a). The main points of interest in the calculation of U_f for $Q(\delta)(\sqrt{\Upsilon})$ using algorithm 3.2 are as follows.

We have $|\sqrt{\gamma}| = \sqrt{9.7} \approx 3.14 < 4$. For $k = 0$ we have $|N_\delta(A_0)| = 1$, $b(\mu_2) \approx 0.8 < 1$ and so algorithm 3.2C is not required. Algorithm 3.1A gives $R_0^{(1)}$ with

$$|N_\delta(A_1^{(1)})| = |N_\delta(A_0)N_\delta(R_0^{(1)})| = \sqrt{17} \neq 4$$

and so $V_f \nearrow A_1$. Midpoint testing is negative and so we reset $k = 1$. Since $|N_\delta(A_1)| > 4$ algorithm 3.2C is once again not required. Algorithm 3.1A gives $R_1^{(1)}$ with

$$|N_\delta(A_2^{(1)})| = |N_\delta(A_1)N_\delta(R_1^{(1)})| = 2 \neq 4$$

and so $V_f \nearrow A_2$. Midpoint testing is positive $(M_2^{(1)}(\sqrt{\gamma}) = (M_2^{(1)}(\sqrt{\gamma}))^*)$ and so $q = 4$ is even. We therefore arrive at step 18 with $d = -2$, $|N_\delta(A_2^{(1)})| = 2$ and so we can take

$$U_f = \delta A_2^{(1)}/2 = ((2-3\delta) - \delta\sqrt{\gamma})/2$$

Note that in spite of the fact that $d = -2$, $|\sqrt{\gamma}| < 4$ the amount of work involved in testing for possibilities (i), (ii), (iii) is trivial.

(b) Let $d = -11$, $\gamma = (3+\delta)/2$ (see example 3.3b). The calculation of U_f using algorithm 3.2 proceeds as follows.

We have $|\sqrt{\gamma}| = \sqrt{3.5} \approx 1.5 < 4$. For $k = 0$ we have $b(\mu_2) \approx 1.67 \geq 1$ and so algorithm 3.2C must be applied. We start with $\beta = 1$. For $\xi = -1$, $s = 1$ steps 7, 8 give $\alpha = (1+\delta)/2$ and we have $N_\delta(A_0)N_\delta((\alpha+\beta W_0)/g_0) = -4$. Furthermore $t \approx 2.94$ and so we set

$$\mu_1 = (1+\delta)/2,\ \mu_2 = 1,\ T = ((1+\delta)/2 + W_0)/g_0,\ mT = 2.94$$

Since $mT \geq |A_0|/2 = 1/2$ we must continue the testing. The next case considered is $\xi = -1$, $s = -1$ and step 8 gives $\alpha = -(1+\delta)/2$ (that is we have found $-T^*$). This case is eventually rejected at step 11 since $t|A_0| = t \approx 1.36 \leq 2$. The remaining cases for $\beta = 1$ (that is $\xi = 1$, $s = \pm 1$) give $\alpha = \pm 5$ and are rejected at step 9. Since no other

$\beta \in Z(\delta)^+ \cap I_0$ satisfies $|\beta| < b(\mu_2)$ we return to step 2.5 of algorithm 3.2 and set

$$U_f = TA_0/2 = ((1+\delta)/2 + \sqrt{\gamma})/2$$

Note that $|V_f| > |A_1|$ (see example 3.3b) but that V_f has been located when $k = 0$. This illustrates the final comments proceeding algorithm 3.2C.

(c) Let $d = -41$, $\gamma = 7 + 4\delta$. $Z(\delta)(\sqrt{\gamma})$ is of form 2 and so we can calculate U_f using algorithm 3.2. Since $d \neq -1, -2, -7$ and $|\sqrt{\gamma}| \geq 4$ algorithm 3.2 will only need to test for the possibility $V_f \sim A_m$. For $k = 17$ we find $N_\delta(A_{k+1}^{(1)}) = 4$ and so we can take

$$U_f = A_{18}^{(1)}/2 = ((31888 - 2305\delta) + (3181 - 942\delta)\sqrt{\gamma})/2$$

(d) Let $d = -47$, $\gamma = 4 + 5\delta$. $Z(\delta)(\sqrt{\gamma})$ is of form 2 and we have $|\sqrt{\gamma}| \geq 4$. Algorithm 3.2 locates the midpoint of a period when $k = 19$. We have $M_{20}^{(1)}(\sqrt{\gamma}) = (M_{20}^{(1)}(\sqrt{\gamma}))^*$ where

$$A_{20}^{(1)} = (-51063 - 4113\delta)/2 + (-4842 + 160\delta)\sqrt{\gamma}$$

We can therefore take

$$\begin{aligned} U_f &= A_{40} \\ &= A_{20}^{(1)}/(A_{20}^{(1)})^* \\ &= (373536196 - 46508043\delta) + (11410054 - 1208007 8\delta)\sqrt{\gamma} \quad // \end{aligned}$$

Although algorithms 3.1, 3.2 allow us to find U_f for a large number of type I quartic fields there are of course many such fields for which these algorithms cannot guarantee to find U_f, that is fields for which $Z(\delta)(\sqrt{\gamma})$ can only be expressed in form 3 or form 4. There are several ways in which we might choose to attack these more difficult cases.

One approach to this problem involves generalizing $M(\sqrt{\gamma})$ and related ideas to the module

$$L(\sqrt{\gamma}) = \{\alpha+(\beta/\gamma_2)\sqrt{\gamma} : \alpha \in Z(\delta),\ \beta \in I(\gamma_1,\gamma_2)\}$$

where $\gamma, \gamma_1, \gamma_2$ are as in theorem 3.2. Note that $L(\sqrt{\gamma}) \subseteq Z(\delta)(\sqrt{\gamma})$ with equality whenever $Z(\delta)(\sqrt{\gamma})$ is of form 3. It is not difficult to see that in general the ideas developed for $M(\sqrt{\gamma})$ in chapter two and this section will extend to $L(\sqrt{\gamma})$. In particular we will find that if $U_f \in L(\sqrt{\gamma})$ then U_f is a relative minimum of $L(\sqrt{\gamma})$ while if $U_f \notin L(\sqrt{\gamma})$ then $V_f = 2U_f \in L(\sqrt{\gamma})$. Consequently generalizations of algorithms 3.1, 3.2 will enable us to calculate U_f. This idea appears to be potentially the most efficient approach to the problem of calculating U_f when $Z(\delta)(\sqrt{\gamma})$ is of form 3 or form 4. However there are a number of practical problems which we have not yet solved. Firstly the search required by the generalized version of algorithm 2.1A is more involved. It appears that it will be necessary to search through the appropriate β coefficients in strict order of increasing magnitude in order to avoid unnecessarily large searches. Secondly the ideals I_k', I_k which arise in the representation of $M_k(\sqrt{\gamma})$ generalize to ideals H_k, J_k in the representation of $L_k(\sqrt{\gamma})$ and we have not as yet been able to establish the precise relationship that exists between H_k, J_k. Consequently the calculation of a representation of $L_k(\sqrt{\gamma})$ is not as simple as we might hope for. Because of the above reasons we shall not pursue this approach in this thesis. However this approach does appear most promising and it is our intention to pursue it at some future point in time. (Note that one of the main reasons for considering $I(\gamma_1,\gamma_2)$ in section one was to prepare the way for this future work).

An alternative approach to the problem of calculating U_f is to use algorithm 3.1 (form 3) or algorithm 3.2 (form 4) to calculate a unit

$U \in Z(\delta)(\sqrt{\gamma})$ and then determine U_f from U using the ideas in Jeans and Hendy [1978] and Jeans [1978].

We shall pursue this alternative approach in the following paragraphs. This approach does have a number of drawbacks from an efficiency point of view which will become evident as we proceed. However the resulting algorithms 3.3, 3.4 are quite satisfactory provided $|U|$ is not too large.

If $\sqrt{-1} \in Q(\Gamma)$ or $\sqrt{-3} \in Q(\Gamma)$ then $Q(\Gamma)$ is of the form $Q(\delta)(\sqrt{\gamma})$ with $d = -1$ or $d = -3$. U_f for such cases can be found using algorithms 3.1, 3.2 since $h(-1) = h(-3) = 1$. Consequently in developing algorithms 3.3, 3.4 we shall assume $\sqrt{-1}, \sqrt{-3} \notin Q(\Gamma)$. The effect of this assumption is to ensure that the only roots of unity in $Q(\Gamma)$ are ± 1 (see theorem 3.9). This assumption simplifies the development of algorithms 3.3, 3.4. Since these two algorithms will be virtually identical we shall develop them both at the same time.

The first step in algorithms 3.3, 3.4 is to check to see if U_f is "small". By this we mean select $L \in \mathbb{R}^+$, $L > 1$ and check to see if $|U_f| \leq L$. (If $U_f = (\alpha+(\beta/\gamma_2)\sqrt{\gamma})/2$, $|U_f| \leq L$ then

$$|\beta| = |\gamma_2||U_f - U_f^*|/|\sqrt{\gamma}| < |\gamma_2|(L+1)/|\sqrt{\gamma}|$$

$$|\alpha-(\beta/\gamma_2)\sqrt{\gamma}| = 2|U_f^*| < 2$$

Thus there are finitely many cases to check. An appropriately modified version of algorithm 3.2C is quite suitable for this purpose). Of course if U_f is found in this search then we can stop. However in general no unit is found and so we have the lower bound

$$|U_f| > L > 1$$

This lower bound is important later in the algorithm but is obtained as

as a first step in order to most simply deal with the cases where $|U_f|$ is "small". The choice of L depends on several factors which we shall discuss at a later point.

The next step is to use algorithm 3.1 (form 3/algorithm 3.3) or algorithm 3.2 (form 4/algorithm 3.4) to calculate a unit $U \in Z(\delta)(\sqrt{\gamma})$. Now

$$U = \xi U_f^m, \quad m \in Z^+, \quad \xi \in \{-1,1\}$$

(recall that we are assuming $\sqrt{-1}, \sqrt{-3} \notin Q(\Gamma)$) and so

$$U_f = (\xi U)^{1/m}$$

Since $|U_f| > L$ we have

$$m < \ell n|U|/\ell n\, L$$

Therefore to determine whether or not U is fundamental we shall check to see if any of the finitely many numbers

$$(\xi U)^{1/p}, \quad p = 2,3,\ldots(p \text{ prime}), \quad p < b(p) = \ell n|U|/\ell n\, L$$

is a unit of $Z(\delta)(\sqrt{\gamma})$. Note that all p p^{th} roots of ξU must be considered when $p \geq 3$. Now if none of these p^{th} roots is a unit of $Z(\delta)(\sqrt{\gamma})$ then U is clearly fundamental. If however $U_1 = (\xi U)^{1/p_1}$ is a unit of $Z(\delta)(\sqrt{\gamma})$ then clearly U is not fundamental. We therefore replace U with U_1, recalculate b(p) and then apply the above testing procedure starting at p_1. Clearly U_f will be obtained in a finite number of repetitions of this procedure.

This completes the outline of algorithms 3.3, 3.4. We now consider the finer details. In particular we consider the choice of L and the method for testing whether or not $(\xi U)^{1/p}$ is a unit of $Z(\delta)(\sqrt{\gamma})$.

The choice of L is influenced by two opposing factors. On the one hand a larger value of L gives a smaller value of b(p) and this reduces the number of roots of ξU which need to be considered. On the other hand a larger value of L involves a larger search in the initial step of algorithms 3.3, 3.4. Note that the best balance between these two factors really depends on the size of $|U|$. Since we shall not encounter any unduly large values of $|U|$ in this thesis we shall simplify matters by using the somewhat arbitrary choice of L = 10.

The testing of the p^{th} roots of ξU can be simplified as follows. Firstly if $(\xi U)^{\frac{1}{2}}$ is a unit of $Z(\delta)(\sqrt{\Upsilon})$ then we must have

$$N_\delta(U) = (N_\delta((\xi U)^{\frac{1}{2}}))^2 = (\pm 1)^2 = 1$$

This provides a very simple initial test for the case p = 2. Secondly for $p \geq 3$ we have

$$(\xi U)^{1/p} = \xi U^{1/p}$$

and so it suffices to consider the case $\xi = 1$. However for p = 2 we must test both $U^{\frac{1}{2}}$ and $(-U)^{\frac{1}{2}} = iU^{\frac{1}{2}}$ whenever $N_\delta(U) = 1$.

It therefore remains to show how we can test whether or not the complex number $(\xi U)^{1/p}$ is a unit of $Z(\delta)(\sqrt{\Upsilon})$.

THEOREM 3.16

Let $Q(\Gamma) = Q(\delta)(\sqrt{\Upsilon})$ be any type I quartic field with $\sqrt{-1}, \sqrt{-3} \notin Q(\Gamma)$ and Υ_2 as in theorem 3.2. Let U be a unit of $Z(\delta)(\sqrt{\Upsilon})$ and $\xi \in \{-1,1\}$. Then

$$U_1 = (\xi U)^{1/p}, \text{ p prime, is a unit of } Z(\delta)(\sqrt{\Upsilon})$$

if and only if

$$\alpha = U_1 + sU_1^{-1} \in Z(\delta),\ \beta = \Upsilon_2(U_1 - sU_1^{-1})/\sqrt{\Upsilon} \in Z(\delta) \tag{10}$$

$$\text{where } s \in \{-1,1\} \text{ with } s = N_\delta(U) \text{ if } p \geq 3$$

PROOF

If U_1 is a unit of $Z(\delta)(\sqrt{\gamma})$ then we can write U_1 using form 4 notation as

$$U_1 = (\alpha+(\beta/\gamma_2)\sqrt{\gamma})/2, \quad \alpha,\beta \in Z(\delta)$$

It is easily checked that

$$\alpha = U_1 + U_1^*, \quad \beta = \gamma_2(U_1 - U_1^*)/\sqrt{\gamma}$$

Since $U_1U_1^* = N_\delta(U_1) = s \in \{-1,1\}$ we have $U_1^* = sU_1^{-1}$. Finally for $p \geq 3$ we have

$$N_\delta(U_1) = (N_\delta(U_1))^p = N_\delta(U_1^p) = N_\delta(\xi U) = N_\delta(U)$$

since p is odd and $N_\delta(\xi) = \xi^2 = 1$.

The reverse implication is trivial to check. //

Thus (10) gives us a method for testing whether or not $(\xi U)^{1/p}$ is a unit of $Z(\delta)(\sqrt{\gamma})$. Of course if $Z(\delta)(\sqrt{\gamma})$ is of form 3 then (10) can be modified to

$$\alpha = (U_1+sU_1^{-1})/2 \in Z(\delta), \quad \beta = \gamma_2(U_1-sU_1^{-1})/2\sqrt{\gamma} \in Z(\delta)$$

Note that when $p = 2$ we do not know the value of s and so both $s = 1$ and $s = -1$ must be tried.

We now collect together the ideas and results presented in the preceding paragraphs to give the required algorithms 3.3, 3.4.

ALGORITHM 3.3

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ be a type I quartic field with $\sqrt{-1}, \sqrt{-3} \notin Q(\Gamma)$ for which $Z(\delta)(\sqrt{\gamma})$ is of form 3. Then a fundamental unit U_f of $Q(\Gamma)$ can be calculated as follows.

1 Given $L > 1$ check to see if $|U_f| \leq L$, if U_f is found during this check then stop

2 Use algorithm 3.1 to calculate a unit $U \in Z(\delta)(\sqrt{\gamma})$

3 Set $b(p) = \ell n|U|/\ell n\, L$

4 If $2 \geq b(p)$ set $U_f = U$ and stop

5 If $N_\delta(U) = -1$ go to 12

6 Calculate $x = \sqrt{U}$, $y = x^{-1}$, set $\xi = 1$, $s = 1$

7 If $\alpha = (x+sy)/2 \notin Z(\delta)$ go to 10

8 If $\beta = \gamma_2(x-sy)/2\sqrt{\gamma} \notin Z(\delta)$ go to 10

9 Reset $U = \alpha+(\beta/\gamma_2)\sqrt{\gamma}$, divide $b(p)$ by 2, go to 4

10 If $s = 1$ set $s = -1$, and go to 7

11 If $\xi = 1$ multiply x by i, set $y = x^{-1}$, $\xi = -1$, $s = 1$, and go to 7

12 Set $p = 3$

13 If $p \geq b(p)$ set $U_f = U$ and stop

14 Set $x = U^{1/p}$ (any p^{th} root of U), $y = N_\delta(U)x^{-1}$

15 Set $\mu = \cos(2\pi/p) + i \sin(2\pi/p)$, $j = 1$

16 If $\alpha = (x+y)/2 \notin Z(\delta)$ go to 19

17 If $\beta = \gamma_2(x-y)/2\sqrt{\gamma} \notin Z(\delta)$ go to 19

18 Reset $U = \alpha+(\beta/\gamma_2)\sqrt{\gamma}$, divide $b(p)$ by p, go to 13

19 If $j < p$ multiply x by μ, set $y = N_\delta(U)x^{-1}$, increment j by 1, and go to 16

20 Increment p to the next largest prime, go to 13 //

ALGORITHM 3.4

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ be a type I quartic field with $\sqrt{-1}, \sqrt{-3} \notin Q(\Gamma)$ for which $Z(\delta)(\sqrt{\gamma})$ is of form 4. Then a fundamental unit U_f of $Q(\Gamma)$

can be calculated using algorithm 3.3 modified as follows.

(a) Replace algorithm 3.1 with algorithm 3.2 in step 2

(b) Delete the divisor 2 from steps 7, 8, 16, 17

(c) Replace the expression for U in steps 9, 18 with

$$U = (\alpha+(\beta/\gamma_2)\sqrt{\gamma})/2$$ //

Steps 10, 11 ensure that we check the four possibilities $\sqrt{U}$, $s = \pm 1$ and $\sqrt{-U}$, $s = \pm 1$ in (10). μ in step 15 is of course a primitive p^{th} root of unity and as j varies from 1 to p the multiplication in step 19 cycles x through the p p^{th} roots of U.

Note that algorithms 3.3, 3.4 as stated implicitly assume that all calculations are exact. Of course this is impossible to carry out in practice and we therefore proceed as follows. Calculations are carried out to sufficient accuracy so as to ensure that α,β in steps 7, 8, 16, 17 are correct to at least 3D. Initial decisions are based on whether or not $\alpha,\beta \in Z(\delta)$ to within rounding error. (That is $cRe(\alpha)$, $cIm(\alpha)/|\delta|$, $cRe(\beta)$, $cIm(\beta)/|\delta| \in Z$ to within rounding error). This approach may fail to reject some cases where the number being tested is not a unit of $Z(\delta)(\sqrt{\gamma})$. Therefore to ensure absolute certainty in the results we must insert a norm test in steps 9 and 18 of algorithms 3.3, 3.4. For example step 9 of algorithm 3.3 becomes

9 If $N_\delta(\alpha+(\beta/\gamma_2)\sqrt{\gamma}) \in \{-1,1\}$ reset $U = \alpha + (\beta/\gamma_2)\sqrt{\gamma}$, divide b(p) by 2, and go to 4

The testing of $U^{1/p}$ requires $O(\ell n|U|/p)$ precision arithmetic. Thus for large $|U|$ and small p the calculations required by algorithms 3.3, 3.4 will involve multiprecision arithmetic. On the other hand when p is large we will have a correspondingly large number of possibilities

to consider although the precision of calculation required does decrease as p increases. A final point to note is that algorithms 3.3, 3.4 can "overshoot" U_f (that is calculate ξU_f^m, $m > 1$) and then backtrack via the root testing procedure to U_f. These are the main drawbacks of algorithms 3.3, 3.4 although they are not particularly significant problems for the cases which we shall encounter in this thesis. (In view of these points it is not difficult to see that the previously mentioned generalization of $M(\sqrt{\Upsilon})$ to $L(\sqrt{\Upsilon})$ (when fully developed) should be much more efficient than algorithms 3.3, 3.4).

We now briefly illustrate the use of algorithms 3.3, 3.4 with several examples.

EXAMPLE 3.5

(a) Let $d = -10$, $\Upsilon = 3 + 2\delta$. Since $N(\Upsilon) = 49$, $7 \nmid \Upsilon$ it is clear from theorem 3.3 that we must take $\Upsilon_2 = \Upsilon$, $\Upsilon_1 = 1$ in theorem 3.2. Thus $Z(\delta)(\sqrt{\Upsilon})$ is of form 3 or form 4. Using theorem 3.5 we find $I(\Upsilon_1, \Upsilon_2) = Z[7, 5+\delta]$. It is then easily checked using theorem 3.6 that $Z(\delta)(\sqrt{\Upsilon})$ is of form 4 but not of form 3. We therefore use algorithm 3.4 to calculate U_f. The search in step 1 ($L = 10$) produces

$$U_f = (1 + ((5+\delta)/\Upsilon_2)\sqrt{\Upsilon})/2$$

and so we stop.

(b) Let $d = -29$, $\Upsilon = 6 + 4\delta$. We have $N(\Upsilon) = 500 = 2^2.5^3$ and so we must take $\Upsilon_2 = \Upsilon$, $\Upsilon_1 = 1$ in theorem 3.2. We have $I(\Upsilon_1, \Upsilon_2) = Z[50, 39+\delta]$ and it then follows from theorem 3.6 that $Z(\delta)(\sqrt{\Upsilon})$ is of form 3. We apply algorithm 3.3 to find U_f.

Step 1 fails to produce U_f and so we conclude $|U_f| > 10$. Step 2 produces the unit

$$U = (-61103449 + 1126323228) + (67292105 + 20419326\delta)\sqrt{\Upsilon}$$

which has $N_\delta(U) = -1$. We set $b(p) = \ell n|U|/\ell n\, 10 \approx 8.8$ and then begin testing the p^{th} roots of ξU, $p < b(p)$. Since $N_\delta(U) = -1$ we see that $U \neq \xi U_f^2$ and so we go to step 12. All three cube roots of U are rejected at step 16 since we find

$$j = 1,\ \alpha \approx 453.393 + 52.441\delta \notin Z(\delta)$$

$$j = 2,\ \alpha \approx -471.266 + 46.693\delta \notin Z(\delta)$$

$$j = 3,\ \alpha \approx 17.873 + 99.134\delta \notin Z(\delta)$$

(In step 14 we have taken x to be the p^{th} root of U for which $0 \leq \arg(U^{1/p}) < 2\pi/p)$. We therefore set $p = 5$ and test the fifth roots of U. For $j = 1$ we find

$$\alpha \approx 31.000 + 2.000\delta \in Z(\delta)$$

$$\beta \approx 86.000 + 24.000\delta \in Z(\delta)$$

A norm calculation confirms that these are indeed the coefficients of a $Z(\delta)(\sqrt{\gamma})$ unit and so we reset

$$U = (31+2\delta) + ((86+24\delta)/\gamma_2)\sqrt{\gamma}$$

We now have $b(p) \approx 8.8/5 \approx 1.8$. The current value of p is 5 and since $5 \geq 1.8$ (step 13) we can take $U_f = U$ and stop. //

We now illustrate the combined use of algorithms 3.1, 3.2, 3.3, 3.4 by calculating U_f for the distinct non-isomorphic quartic fields $Q(\sqrt[4]{D})$ where D is a negative integer lying in the range

$$-99 \leq D \leq -1 \qquad (11)$$

THEOREM 3.17

Any quartic field formed by adjoining the fourth root of a negative rational number to Q is a type I quartic field of the form

$$Q(\sqrt[4]{D}) = Q(\delta)(\sqrt{\gamma})$$

$$D = rs^2t^3,\ -r,s,t \in Z^+,\ d = rt,\ \delta = \sqrt{d},\ \gamma = st\delta = \sqrt{D}$$

with r,s,t square-free and relatively prime in pairs and $D \neq -4$. γ is a non-square rational square-free $Z(\delta)$ integer.

Let $D_1 = (-t)s^2|r|^3$ and $\kappa = s|r|\delta$. Then $Q(\delta)(\sqrt{\gamma})$ and $Q(\delta)(\sqrt{\kappa})$ are conjugate fields.

PROOF

The fact that the field will be of the form $Q(\sqrt[4]{D})$ with D a fourth power free negative rational integer is trivial. We can clearly represent such a D in the form stated and $d = rt$ is clearly a square-free negative rational integer. If $\gamma = st\delta = \alpha^2, \alpha \in Z(\delta)$ then $(N(\alpha))^2 = |r|s^2t^3$. This implies $rt = -1$, that is

$$\gamma = s\sqrt{-1} = (a+b\sqrt{-1})^2 = a^2 - b^2 + 2ab\sqrt{-1},\ a,b \in Z$$

Thus $|a| = |b| = 1$ (s is squarefree) and $D = -4 = (\pm 1 \pm \sqrt{-1})^4$. Note that $Q(\sqrt[4]{-4}) = Q(\sqrt{-1})$ is a quadratic field. Therefore if $D \neq -4$ we have that γ is a non-square rational squarefree $Z(\delta)$ integer and so $Q(\sqrt[4]{D}) = Q(\delta)(\sqrt{\gamma})$ is indeed a type I quartic field.

Note that $\kappa'\gamma = (rst)^2$ and so it follows that

$$Q(\delta)(\sqrt{\gamma}) = Q(\delta)(\sqrt{\kappa'})$$

(See the comments following theorem 3.1). The final result of the theorem is now clear. //

We can therefore confine our attention to the fields

$$Q(\sqrt[4]{D}),\ D = rs^2t^3 \text{ as in theorem 3.17},\ |r| \geq t \qquad (12)$$

since theorem 3.17 shows that any other quartic field of the form $Q(\sqrt[4]{a})$,

$a \in Q$, $a < 0$ is either equal to or conjugate to one of the fields in (12). The next result is used to decide which algorithm should be used to calculate U_f for a given $Q(\sqrt[4]{D})$.

THEOREM 3.18

Let $Q(\sqrt[4]{D}) = Q(\delta)(\sqrt{\gamma})$ be as described in theorem 3.17.

(a) In theorem 3.2 we can take

$$\gamma_1 = s\delta/(s,\Delta),\ \gamma_2 = (s,\Delta)t$$

(Δ is the discriminant of $Q(\delta)$). If

$$t = (s,\Delta) = 1 \tag{13}$$

then $Z(\delta)(\sqrt{\gamma})$ is of form 1 or form 2. (Note $(s,d) = (s,rt) = 1$ and so $(s,\Delta) \mid 2$).

(b) If

$$\left.\begin{array}{l} \text{either } rt \equiv 1 \pmod 8 \text{ and } st \equiv 2 \pmod 4 \\ \text{or } rt \equiv 5 \pmod 8 \text{ and } st \equiv 1,2,3 \pmod 4 \\ \text{or } rt \equiv 2,3 \pmod 4 \text{ and } st \equiv 1,3 \pmod 4 \end{array}\right\} \tag{14}$$

then $Z(\delta)(\sqrt{\gamma})$ is of form 1 or form 3.

(c) $Z(\delta)(\sqrt{\gamma})$ is of form 1

if and only if

r,s,t satisfy (13) and (14)

(d) If $\sqrt{-1} \in Q(\sqrt[4]{D})$ then $d = rt = -1$, $D = -s^2$, $\gamma = s\sqrt{-1}$.

(i) If s is odd then γ is square free in $Z(\sqrt{-1})$.

(ii) If s is even then $\gamma = (1+\sqrt{-1})^2(s/2)$. We have

$$Q(\sqrt[4]{-s^2}) = Q(\sqrt{-1})(\sqrt{s/2})$$

with $s/2$ square free in $Z(\sqrt{-1})$ and $Z(\sqrt{-1})(\sqrt{s/2})$ is of form 2.

(e) If $\sqrt{-3} \in Q(\sqrt[3]{D})$ then $d = rt = -3$. If $|r| \geq t$ (that is $r = -3$, $t = 1$) then $\gamma = s\sqrt{-3}$ is square free in $Z(\sqrt{-3})$. (Note that the case $\sqrt{-1}, \sqrt{-3} \in Q(\sqrt[3]{D})$ occurs when $D = -36$).

PROOF

(a) The choice of γ_1, γ_2 is clear in view of theorem 3.3. Since (13) implies $\gamma_2 = 1$ the second result is also clear.

(b) This result follows from theorem 3.7. Note that when $d \equiv 1 \pmod 4$ we have $2\delta \equiv 2 \pmod 4$ and $3\delta \equiv 2 + \delta \pmod 4$.

(c) This result follows from theorem 3.8. (γ_1 in (a) is best possible and (14) lists all cases satisfying one of the congruences in theorem 3.7).

(d) Suppose $d \neq -1$. Then $Q(\delta)(\sqrt{\gamma}) = Q(\delta)(\sqrt{-1})$ and it follows that $(-1)\gamma = \alpha^2, \alpha \in Z(\delta)$. (See comments following theorem 3.1). Thus $(N(\alpha))^2 = |r|s^2t^3$ and so $|r|t = 1$, that is $d = -1$. This is a contradiction and so the first result of the theorem does indeed hold. Parts (i), (ii) are easily checked with the aid of theorems 3.3, 3.7.

(e) This case can be proved in a similar manner to part (d). //

This result is not best possible in that it may not recognize the simplest form for $Z(\delta)(\sqrt{\gamma})$. However a better result would be both tedious to derive and messy to present. Theorem 3.18 will suffice for our purposes. We therefore take the form of $Z(\delta)(\sqrt{\gamma})$, $\delta = \sqrt{rt}$, $\gamma = st\delta$ to be as follows.

$$\left.\begin{array}{ll} \text{form 1} & r,s,t \text{ satisfy (13), (14)} \\ \text{form 2} & r,s,t \text{ satisfy (13)} \\ \text{form 3} & r,s,t \text{ satisfy (14)} \\ \text{form 4} & \text{all remaining cases} \end{array}\right\} \qquad (15)$$

U_f for $Q(\sqrt[3]{D})$, D as in (12), can therefore be calculated as follows.

(i) If $D \neq -s^2$ or $2 \nmid s$ then use algorithm 3.j applied to $M(\sqrt{\gamma})$, $d = rt$ where j is the form of $Z(\delta)(\sqrt{\gamma})$ determined by (15).

(ii) If $D = -s^2$, $2|s$ then use algorithm 3.2 applied to $M(\sqrt{s/2})$, $d = -1$.

(Recall that if $\sqrt{-1} \in Q(\Gamma)$ or $\sqrt{-3} \in Q(\Gamma)$ then we cannot use algorithms 3.3, 3.4. The condition $|r| \geq t$ in (12) and the special case for $D = -s^2$, $2|s$ ensure that only algorithms 3.1, 3.2 are used when $\sqrt{-1} \in Q(\sqrt[4]{D})$ or $\sqrt{-3} \in Q(\sqrt[4]{D})$).

A fundamental unit for each D satisfying both (11) and (12) is given in table 3.3. (All other values of D satisfying (11) have a self-explanatory note). The entries in table 3.3 are as follows. D, d, st are as in theorem 3.17. The entry j in the form column indicates that (15) determines $Z(\delta)(\sqrt{\gamma})$ as being of form j and that algorithm 3.j was used to calculate U_f. The entry * in the form column indicates that $D = -s^2$, $2|s$ and that U_f has been calculated as in (ii) above. The rational integers a, b, e, f are the coefficients of two $Z(\delta)$ integers

$$\alpha = (a+b\delta)/c, \quad \beta = (e+f\delta)/c$$

and α, β are the coefficients of a $Z(\sqrt[4]{D})$ integer

$$A = \begin{cases} \alpha + \beta\sqrt{\gamma} & \text{form 1} \\ (\alpha + \beta\sqrt{\gamma})/2 & \text{form 2} \\ \alpha + (\beta/\gamma_2)\sqrt{\gamma} & \text{form 3} \\ (\alpha + (\beta/\gamma_2)\sqrt{\gamma})/2 & \text{form 4} \\ (\alpha + \beta\sqrt{s/2})/2 & \text{form *} \end{cases}$$

where $\gamma_2 = (s,\Delta)t$. We have

$$U_f = \begin{cases} A & \text{if } |N_\delta(A)| = 1 \\ A/A^* & \text{if } |N_\delta(A)| > 1 \end{cases}$$

TABLE 3.3 U_f for $Q(\sqrt{D})$, $-99 \le D \le -1$

D	d	st	form	a	b	e	f	$N_\delta(A)$
-1	-1	1	1	-1	-1	-1	0	i
-2	-2	1	1	-1	1	0	1	-1
-3	-3	1	1	-1	1	1	1	1
-4			not a quartic field					
-5	-5	1	1	2	1	2	0	-1
-6	-6	1	1	1	-4	-4	-2	1
-7	-7	1	2	3	-1	-1	-1	1
-8			see D=-2 (conjugate field)					
-9	-1	3	1	1	1	1	0	-i
-10	-10	1	1	27	1	12	-3	-1
-11	-11	1	1	-9	5	3	3	1
-12	-3	2	1	-2	4	2	2	1
-13	-13	1	1	-86	3	-28	10	-1
-14	-14	1	1	-13	2	-2	2	1
-15	-15	1	2	14	-2	2	-2	1
-16			see D=-1					
-17	-17	1	1	-16	-1	-7	1	1
-18	-2	3	1	-3	-2	-2	0	1
-19	-19	1	1	8115	-1395	689	-1103	1
-20	-5	2	4	1	1	2	0	-1
-21	-21	1	1	1	-24	-36	-8	1
-22	-22	1	1	-91167	-267972	-440140	-81146	1
-23	-23	1	2	19	7	17	1	1
-24	-6	2	4	-190	-40	-184	24	1
-25	-1	5	1	0	-3	-1	-1	1
-26	-26	1	1	-125	17	-12	13	-1
-27			see D=-3 (conjugate field)					
-28	-7	2	1	-12	2	-2	2	1
-29	-29	1	1	-330206	-189709	-411912	-39122	-1
-30	-30	1	1	-74879	-21012	-57396	-2218	1
-31	-31	1	2	-4398	338	-754	338	1
-32			see D=-2					
-33	-33	1	1	67	32	74	6	1
-34	-34	1	1	-33	8	4	4	1
-35	-35	1	1	527	65	265	-7	1
-36	-1	6	*	1	-1	1	-1	i
-37	-37	1	1	-571878	-71847	-289258	6356	-1
-38	-38	1	1	24667022	-2994169	1768526	-1992368	-2
-39	-39	1	2	-35	5	-1	3	1
-40	-10	2	4	-26	6	-4	8	-1
-41	-41	1	1	-155	8	-29	9	-1
-42	-42	1	1	1	-2700	-4860	-750	1
-43	-43	1	1	-8082321	-1512339	-4970221	-77261	1
-44	-11	2	1	19010	-20340	-13302	-7158	1
-45	-5	3	1	6	-1	1	-1	1
-46	-46	1	1	-45	-14	-38	-2	1
-47	-47	1	2	2161346	-4153806	-7106826	-1206918	1
-48			see D=-3					
-49	-1	7	1	-2	-2	-1	0	i
-50	-2	5	1	-23	-5	-8	3	-1

Cont'd...

TABLE 3.3 (cont'd)

D	d	st	form	a	b	e	f	$N_\delta(A)$
-51	-51	1	1	767	217	613	29	1
-52	-13	2	4	3	-3	-4	-2	-1
-53	-53	1	1	-114662	-319	-30658	4044	-1
-54			see D=-24 (conjugate field)					
-55	-55	1	2	51	3	19	-1	1
-56	-14	2	4	2	16	32	8	1
-57	-57	1	1	-3647	2408	3740	744	1
-58	-58	1	1	621	23	204	-15	-1
-59	-59	1	1	-1674315	-642821	-1686939	-108393	-2
-60	-15	2	1	-8	-2	-4	0	1
-61	-61	1	1	-116682	9525	-10700	6190	-1
-62	-62	1	1	1	16	32	4	1
-63	-7	3	2	-32	-12	-16	0	1
-64			not a quartic field					
-65	-65	1	1	8	1	4	0	-1
-66	-66	1	1	1	-16	-32	-4	1
-67	-67	1	1	1771721	-208521	16041	-105033	-2
-68	-17	2	4	8	2	8	0	-1
-69	-69	1	1	5343	-4790	-8451	-1333	3
-70	-70	1	1	-733471199	-111894660	-408164940	-5922834	1
-71	-71	1	2	-67	-127	-277	-29	1
-72			see D=-18 (conjugate field)					
-73	-73	1	1	-127749	-28800	-90430	-3350	1
-74	-74	1	1	475	-7	100	-15	-1
-75	-3	5	1	-13	5	-1	3	1
-76	-19	2	1	-4254	1020	46	478	1
-77	-77	1	1	-342	49	21	21	1
-78	-78	1	1	-5313	-60689	-128796	-14297	3
-79	-79	1	2	-225146974	18289410	-14844424	10345896	†
-80			see D=-5					
-81			see D=-1					
-82	-82	1	1	-647	-297	-784	-53	-1
-83	-83	1	1	-172243	41567	48365	14167	-2
-84	-21	2	4	170	-10	58	-22	1
-85	-85	1	1	-8549942	-96399	-2198074	193516	-1
-86	-86	1	1	-75708867722	-10607858429	-40421693626	-567485420	-2
-87	-87	1	2	62	-2	10	-2	1
-88	-22	2	4	-15145854	10349976	15421928	6269960	1
-89	-89	1	1	-510642	32360	-47277	19911	-2
-90	-10	3	1	-79	36	8	14	1
-91	-91	1	1	81083	33777	92331	5787	1
-92	-23	2	1	-44	-526	-586	-118	1
-93	-93	1	1	2480919	-155458	223542	-93976	-3
-94	-94	1	1	-9117	2	-2066	214	1
-95	-95	1	2	-84961916	26324512	38870144	7936624	1
-96			see D=-6					
-97	-97	1	1	9172224	-1153609	-493331	-469763	1
-98	-2	7	1	-97	56	-4	28	1
-99	-11	3	1	35	15	19	1	1

†$(-7+\delta)/2$

If $|N_\delta(A)| > 1$ then A is the relative minimum at the midpoint of an even length period of $M(\sqrt{\gamma})$ and $U_f = A_q$. We have only resorted to this format in those cases where the coefficients of U_f are too large for the table.

This completes our own work on units of type I quartic fields. We finish off this section by noting some of the existing literature which is relevant to this area of work. We begin with several results which are specific to certain subtypes of the type I quartic fields.

In chapter two section five we noted two such specific methods, that is the $Z(\delta)CF$ algorithm and Amara's [1981] algorithm. Amara's algorithm applies to type I quartic fields of the form $Q(\delta)(\sqrt{\gamma})$ where $h(d) = 1$. The algorithm is guaranteed to find U_f in all such cases although some of the more practical details are only developed for Euclidean $Q(\delta)$. The $Z(\delta)CF$ algorithm (and variations) can be used to calculate units of type I quartic fields $Q(\delta)(\sqrt{\gamma})$ when $Q(\delta)$ is Euclidean. Lakein [1971, 1974, 1975] shows that $Z(\delta)CF$ algorithms are remarkably successful in calculating U_f when $d = -1, -3$ but notes that there is little theoretical work to back up these observations. (The exception being the work by Stein [1927] on the algorithm of J. Hurwitz [1902] which applies when $d = -1$ and is guaranteed to find at worst U_f^3). Recall that algorithms 3.1, 3.2 are guaranteed to find U_f for any type I quartic field $Q(\delta)(\sqrt{\gamma})$ for which $h(d) = 1$ and so the performance of our algorithms compares most favourably with the $Z(\delta)CF$ algorithm and Amara's algorithm. Furthermore the application of our algorithms to the problem of finding fundamental units extends far beyond the case $h(d) = 1$.

We now note a result from the literature which applies to type I quartic fields of the form $Q(\Gamma) = Q(\delta)(\sqrt{d_1})$ where d_1 is a square free rational integer. Note that $Q(\Gamma)$ has three quadratic subfields $Q(\delta)$, $Q(\sqrt{d_1})$, and $Q(\sqrt{d_2})$ where $d_2 = dd_1/(d,d_1)^2$ is a square free integer.

Since $d < 0$ we see that d_1, d_2 are of opposite sign. Thus $Q(\Gamma)$ has one real and two complex quadratic subfields. Without loss of generality we can assume $d_1 > 0$, $d_2 < 0$, and $|d| < |d_2|$. Results (a), (b), (c) below are basically given in Buell et al [1977] although they credit Kuroda [1943] and Kubota [1953] as being largely responsible for results (a), (b). (In the following results $\varepsilon(d_1)$, δ_1, ω_1 are the obvious integers of $Q(\sqrt{d_1})$ and U_f as usual denotes a fundamental unit of $Q(\Gamma)$ satisfying $|U_f| > 1$).

(a) If $N(\varepsilon(d_1)) = -1$ then we can take $U_f = \varepsilon(d_1)$

(b) If $N(\varepsilon(d_1)) = 1$ then we can take

$$U_f = \begin{cases} \sqrt{\xi\varepsilon(d_1)} & \text{if } g\varepsilon(d_1) = \alpha^2,\ \alpha \in Z(\delta_1) \\ \varepsilon(d_1) & \text{otherwise} \end{cases}$$

where

$$\xi = \begin{cases} i & \text{if } d = -1 \\ -1 & \text{otherwise} \end{cases}, \qquad g = \begin{cases} 2 & \text{if } d = -1 \\ -d & \text{otherwise} \end{cases}$$

If α exists then we have

$$\sqrt{\xi\varepsilon(d_1)} = \begin{cases} \pm(1+i)\alpha/2 & \text{if } d = -1 \\ \pm\alpha/\delta & \text{otherwise} \end{cases}$$

(c) The existence of α in (b) can be determined from the simple continued fraction expansion of $\omega_1 = (a_1, \overline{a_2, \ldots, a_{r+1}})$. (See chapter one section three). We have

α exists

if and only if

$$N(p_{k-1} - q_{k-1}\omega_1) = \pm g, \quad k-1 = r/2$$

in which case we can take $\alpha = p_{k-1} - q_{k-1}\omega_1'$.

Thus to determine U_f we expand ω_1 using algorithm 1.2 until the midpoint of a period is reached. If at this midpoint we have $P_{k+1} = P_k$ (even length period) and $N(p_{k-1}-q_{k-1}\omega_1) = \pm g$, that is $Q_k/Q_1 = g$, then we can take $\alpha = p_{k-1}-q_{k-1}\omega_1'$ and

$$U_f = \begin{cases} (1+i)\alpha/2 & d = -1 \\ \alpha/\delta & \text{otherwise} \end{cases}$$

In all other cases we can take $U_f = \varepsilon(d_1)$.

Although our algorithms can be used to calculate U_f for this type of field it is clear that the above method is far more efficient.

To finish this brief look at the case $Q(\delta)(\sqrt{d_1})$ we note two final points. Firstly lemma 6 of Buell et al [1977] (when corrected) gives the result that if $(d,d_2) > 2$ or $(d,d_2) = 2$, $d_1 \not\equiv 3 \pmod 4$ then α in (b) above does not exist, that is we can take $U_f = \varepsilon(d_1)$. (Lemma 6 wrongly excludes the case $(d,d_2) = 2$, $d_1 \equiv 3 \pmod 4$). Secondly Nagell [1961, p361] notes that $2\varepsilon(d_1) = \alpha^2$, $\alpha \in Z(\delta_1)$ is impossible when $d_1 \equiv 1 \pmod 4$ since $\langle 2 \rangle$ is not the square of a $Z(\delta_1)$ ideal. (See theorem 1.3). This result clearly extends to the case $2|d$, $d_1 \equiv 1 \pmod 4$. That is $g\varepsilon(d_1) = -d\varepsilon(d_1) = \alpha^2$ is impossible when $2|d$, $d_1 \equiv 1 \pmod 4$. Thus if $d_1 \equiv 1 \pmod 4$ and either $d = -1$ or $2|d$ then we can take $U_f = \varepsilon(d_1)$.

Results and methods of a more general nature have also been applied to the problem of calculating fundamental units of certain type I quartic fields. We note two such methods. Szekeres has indicated (verbally) that his algorithm (Szekeres [1970]) has been applied to the problem of calculating units of fields of the form $Q(\sqrt[4]{D})$, $-D \in Z^+$. However no details are as yet available. Secondly Shanks [1977] in a brief note concerning a table of Ljunggren [1934] indicates that he is able to calculate fundamental units of fields of the form $Q(\sqrt[4]{D})$, $D \in Z$,

(and also other unspecified fields) by making use of the fact that the corresponding "Dedekind zeta functions are expressible in terms of Epstein zeta functions". Further details are not available. Whether or not these methods are applicable to more general type I quartic fields is not clear.

Finally we note that the literature contains a number of algorithms which are designed to calculate units of arbitrary degree fields and which could therefore be applied to the problem of calculating fundamental units of type I quartic fields. (See also the comments in the first few paragraphs of chapter two section five). However in spite of their apparently general nature we often find that these methods have had their main application in cubic fields and their value in fields of higher degree (type I quartic fields in particular) is generally not at all clear. Note also that such algorithms generally suffer the drawback of not being able to take advantage of the specific properties of the various types of fields to which they are applied.

To conclude we note that our algorithms appear to be as successful as any in calculating fundamental units of type I quartic fields. However there is of course room for much improvement especially for the form 3 and form 4 cases. The suggested generalization of $M(\sqrt{Y})$ noted earlier in this section would be a significant step in this direction and would result in a far more uniform and efficient approach to the problem of calculating fundamental units of type I quartic fields.

SECTION THREE

UNITS OF TYPE IIb QUARTIC FIELDS

We finish this chapter by noting a connection between type IIb quartic fields and type I quartic fields which enables us to use the results of the previous section to calculate fundamental systems of units for type IIb quartic fields. This connection is a generalization of the special case noted by Parry [1980] which relates the type IIb quartic field $Q(\sqrt[4]{D})$ to the type I quartic field $Q(\sqrt[4]{-4D})$ where D is a fourth power-free non-square positive rational integer.

We begin with some notation.

DEFINITION 3.3

Let $Q(\delta)(\sqrt{\Upsilon})$ be any type I or type IIb quartic field. For $A = \alpha + \beta\sqrt{\Upsilon} \in Q(\delta)(\sqrt{\Upsilon})$ we define

$$A^* = \alpha - \beta\sqrt{\Upsilon}$$

$$A' = \alpha' + \beta'\sqrt{\Upsilon'}$$

$$A^{*\prime} = A'^* = \alpha' - \beta'\sqrt{\Upsilon'}$$ //

A^*, A', A'^* are of course the conjugates of A with respect to the extension $Q(\delta)(\sqrt{\Upsilon})$ of Q.

Throughout the rest of this section we shall assume that $Q(\Gamma) = Q(\delta)(\sqrt{\Upsilon})$ is a type IIb quartic field. (We shall use $Q(\delta_1)(\sqrt{\kappa})$ to denote the relevant type I quartic field). Thus $Q(\delta)$ is a real quadratic field and Υ is a non-square rational square-free $Z(\delta)$ integer for which Υ, Υ' are of opposite sign. Note that $Q(\delta)(\sqrt{\Upsilon})$, $Q(\delta)(\sqrt{\Upsilon'})$ are isomorphic fields. Therefore without loss of generality we can assume $\Upsilon > 0$, that is $Q(\delta)(\sqrt{\Upsilon})$ is a real field. Note that for $A = \alpha + \beta\sqrt{\Upsilon} \in Q(\delta)(\sqrt{\Upsilon})$ we therefore have

$$\overline{A'} = \overline{\alpha' + \beta'\sqrt{\gamma'}} = \alpha' - \beta'\sqrt{\gamma'} = A^{*'} \tag{16}$$

since $\alpha', \beta' \in \mathbb{R}$ and $\sqrt{\gamma'} = |\sqrt{\gamma'}|i$.

In theorem 3.11 we noted the basic structure of the unit group of $Z(\delta)(\sqrt{\gamma})$. The following theorem which is basically taken from Nagell [1961] gives a more detailed result.

THEOREM 3.19

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$, $\gamma > 0$, be a type IIb quartic field. Then $Z(\delta)(\sqrt{\gamma})$ contains units satisfying

$$UU^* = 1 \tag{17}$$

Such units also satisfy

$$U'U'^* = 1, \quad |U'| = |U'^*| = 1 \tag{18}$$

There is a unique unit $U_1 \in Z(\delta)(\sqrt{\gamma})$ satisfying (17) which also satisfies $U_1 > 1$, U_1 minimal. The units of $Z(\delta)(\sqrt{\gamma})$ which satisfy (17) are precisely those units of the form

$$U = \pm U_1^k, \quad k \in Z$$

A fundamental system of units for $Z(\delta)(\sqrt{\gamma})$ is given by

either $\{\sqrt{\varepsilon(d)}, U_1\}$ if $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ (that is $Q(\Gamma) = Q(\sqrt{\varepsilon(d)})$)

or $\{\varepsilon(d), \sqrt{\varepsilon(d)U_1}\}$ if $\sqrt{\varepsilon(d)U_1} \in Q(\Gamma)$

or $\{\varepsilon(d), U_1\}$ otherwise

($\varepsilon(d)$ is the fundamental unit of $Q(\delta)$).

PROOF

With the exception of (18) these results are given by Nagell [1961, section 2.12]. The results in (18) are simple consequences of (16) and (17). //

We now give the main result of this section which shows that the unit U_1 in theorem 3.19 can be calculated from a fundamental unit of a certain type I quartic field. This result is the generalization of the result in Parry [1980] which was noted at the beginning of this section.

THEOREM 3.20

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$, $\gamma = (a+b\delta)/c > 0$, be a type IIb quartic field with U_1 the unit described in theorem 3.19. Let

$$N(\gamma) = m^2 d_1,\ m \in Z^+,\ d_1 \in Z,\ d_1 \text{ square-free},\ \delta_1 = \sqrt{d_1}$$

Then $Q(\delta_1)$ is a complex quadratic field. Let

$$\kappa = (2a/c + 2m\delta_1)/\ell^2,\ \ell = \begin{cases} 2 \text{ if } 4 \mid (2a/c + 2m\delta_1) \\ 1 \text{ otherwise} \end{cases}$$

Then κ is a non-square rational square-free $Z(\delta_1)$ integer. Let U_f, $|U_f| > 1$, be a fundamental unit of the type I quartic field $Q(\delta_1)(\sqrt{\kappa})$. Then

$$U_f U_f' = U_f \bar{U}_f = |U_f|^2 = U_1^n,\ n \in \{1,2\}$$

PROOF

Since $N(\gamma) < 0$ we have $d_1 < 0$ and so $Q(\delta_1)$ is indeed a complex quadratic field. Since $c|2$ it is clear that κ is a $Z(\delta_1)$ integer. Note that $\sqrt{\kappa'} = \sqrt{\bar{\kappa}} = \overline{\sqrt{\kappa}}$. By squaring it is easily checked that

$$\ell\sqrt{\kappa} = \sqrt{\gamma} + \sqrt{\gamma'},\ \ell\sqrt{\kappa'} = \sqrt{\gamma} - \sqrt{\gamma'} \qquad (19)$$

Using (19) it is not difficult to check that κ is rational square-free. (If $p^2|\kappa$ then $p^2|\gamma$ which contradicts the fact that γ is assumed to be rational square-free). Furthermore if $\sqrt{\kappa} \in Q(\delta_1)$ then $\sqrt{\kappa'} \in Q(\delta_1)$ and so from (19) we have $\sqrt{\gamma} \in Q(\delta_1)$. This implies $Q(\delta)(\sqrt{\gamma}) = Q(\delta)(\delta_1)$ which contradicts the fact that a type IIb quartic field does not have a complex quadratic subfield. Thus κ is non-square and rational

square-free and $Q(\delta_1)(\sqrt{\kappa})$ is a type I quartic field. Note that $B \in Q(\delta_1)(\sqrt{\kappa})$ implies $B' = \overline{B}$.

Suppose $A = \alpha + \beta\sqrt{\gamma} \in Q(\delta)(\sqrt{\gamma})$. Then $B = AA'$ satisfies

$$B = (\alpha+\beta\sqrt{\gamma})(\alpha'+\beta'\sqrt{\gamma'})$$

$$= N(\alpha) + N(\beta)\sqrt{N(\gamma)} + \alpha\beta'\sqrt{\gamma'} + \alpha'\beta\sqrt{\gamma}$$

$$= N(\alpha) + N(\beta)m\delta_1 + \frac{(\alpha\beta'+\alpha'\beta)}{2}(\sqrt{\gamma} + \sqrt{\gamma'})$$

$$- \frac{(\alpha\beta'-\alpha'\beta)}{2}(\sqrt{\gamma} - \sqrt{\gamma'})$$

$$= N(\alpha) + N(\beta)m\delta_1 + \frac{(\alpha\beta'+\alpha'\beta)}{2}\ell\sqrt{\kappa} - \frac{(\alpha\beta'-\alpha'\beta)}{2}\ell\sqrt{\kappa'}$$

$$= N(\alpha) + N(\beta)m\delta_1 + \left(\frac{(\alpha\beta'+\alpha'\beta)}{2} - \frac{(\alpha\beta'-\alpha'\beta)}{2\kappa}\sqrt{N(\kappa)}\right)\ell\sqrt{\kappa}$$

Now $N(\alpha), N(\beta), \alpha\beta' + \alpha'\beta \in Q$, and

$$\alpha\beta' - \alpha'\beta = e\delta,\ e \in Q,\ \sqrt{N(\kappa)} = 2|b|\delta/c\ell^2$$

Thus $(\alpha\beta' - \alpha'\beta)\sqrt{N(\kappa)} \in Q$ and it is now clear that

$$B = \theta + \lambda\sqrt{\kappa},\ \theta,\lambda \in Q(\delta_1)$$

that is $B \in Q(\delta_1)(\sqrt{\kappa})$. Furthermore it is not difficult to check that

$$B^* = \theta - \lambda\sqrt{\kappa} = (\alpha-\beta\sqrt{\gamma})(\alpha'-\beta'\sqrt{\gamma'}) = A^*A'^*$$

(This result is not quite as trivial as it may appear since B, A, A' belong to distinct quartic fields).

Now suppose $B = \theta + \lambda\sqrt{\kappa} \in Q(\delta_1)(\sqrt{\kappa})$. Then in a similar manner to the above we have $C = BB' = B\overline{B} = |B|^2$ satisfies

$$C = N(\theta) + N(\lambda)\sqrt{N(\kappa)} + ((\theta\lambda'+\theta'\lambda) - (\theta\lambda'-\theta'\lambda)\sqrt{N(\gamma)}/\gamma)\sqrt{\gamma}/\ell$$

$$= \eta + \tau\sqrt{\gamma},\quad \eta,\tau \in Q(\delta)$$

that is $C \in Q(\delta)(\sqrt{\Upsilon})$. Once again it is easily checked that

$$C^* = \eta - \tau\sqrt{\Upsilon} = (\theta-\lambda\sqrt{\kappa})(\theta'-\lambda'\sqrt{\kappa'}) = B^*B'^*$$

We now use the results of the previous two paragraphs to obtain the main result of the theorem. We have

$$U = U_f U_f' = U_f \bar{U}_f = |U_f|^2 \in Q(\delta)(\sqrt{\Upsilon})$$

Since U_f is a unit of $Q(\delta_1)(\sqrt{\kappa})$ and U_f' is a unit of $Q(\delta_1)(\sqrt{\kappa'})$ it follows that U is a unit of $Q(\delta)(\sqrt{\Upsilon})$. Note that

$$UU^* = U_f U_f' U_f^* U_f'^* = N_{\delta_1}(U_f)(N_{\delta_1}(U_f))' = 1$$

and that $U = |U_f|^2 > 1$. Therefore by theorem 3.19 we have $U = U_1^n$, $n \in Z^+$. We also have that U_1U_1' is a unit of $Q(\delta_1)(\sqrt{\kappa})$ and so $U_1U_1' = \xi U_f^k$, $k \in Z$, ξ a root of unity. Therefore

$$UU' = (U_1U_1')^n = \xi^n U_f^{kn}$$

Since $|U'| = 1$ (see theorem 3.19) and $U = |U_f|^2$ we have $kn = 2$, that is $n \in \{1,2\}$. //

The basic procedure for finding a fundamental system of units of $Q(\Gamma)$ is therefore as follows.

ALGORITHM 3.5

A fundamental system of units for the type IIb quartic field $Q(\Gamma) = Q(\delta)(\sqrt{\Upsilon})$, $\Upsilon = (a+b\delta)/c > 0$ can be found as follows.

1 Calculate $\varepsilon(d)$ using the simple continued fraction algorithm (see chapter one section three)

2 Set $\kappa = (2a/c+2m\delta_1)/\ell^2$ as in theorem 3.20

3 Use the appropriate algorithm of section two to calculate U_f for the type I quartic field $Q(\delta_1)(\sqrt{\kappa})$

4 Calculate U_1 for $Q(\Gamma)$ (see algorithm 3.5A below)

5 If $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ (see algorithm 3.5B below) then a fundamental system of units for $Q(\Gamma)$ is $\{\sqrt{\varepsilon(d)}, U_1\}$ and we stop

6 If $\sqrt{\varepsilon(d)U_1} \in Q(\Gamma)$ (see algorithm 3.5C below) then a fundamental system of units for $Q(\Gamma)$ is $\{\varepsilon(d), \sqrt{\varepsilon(d)U_1}\}$ and we stop

7 A fundamental system of units for $Q(\Gamma)$ is $\{\varepsilon(d), U_1\}$

8 Stop //

It therefore remains to explain steps 4, 5, 6 in greater detail. To simplify the presentation of the necessary results we shall normally write integers of $Q(\delta)(\sqrt{\gamma})$ in the form

$$(\alpha+\beta/\sqrt{\gamma})/2,\ \alpha \in Z(\delta),\ \beta \in I(1,\gamma),\ \alpha^2 \equiv \beta^2/\gamma \pmod 4$$

That is we take $\gamma_1 = 1, \gamma_2 = \gamma$ in theorem 3.2 and simplify $(\beta/\gamma_2)\sqrt{\gamma}$ to $\beta/\sqrt{\gamma}$. However in some cases where $2|\alpha$, $2|\beta$ we shall cancel the factor 2. The fact that $Z(\delta)(\sqrt{\gamma})$ may be of one of the simpler forms given in definition 3.2 is irrelevant as far as the calculations required by steps 4, 5, 6 are concerned. In fact we shall not even need to know $I(1,\gamma)$.

The ideas presented below are much the same as those used in algorithms 3.3, 3.4. We begin with a result which basically corresponds to theorem 3.16. This result is used in each of steps 4, 5, 6.

THEOREM 3.21

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$, $\gamma > 0$, be a type IIb quartic field with $A = (\alpha+\beta/\sqrt{\gamma})/2 \in Z(\delta)(\sqrt{\gamma})$, $\alpha = (e+f\delta)/c$, $\beta = (m+n\delta)/c \in Z(\delta)$. Then

$$e = c(A+A^*+A'+\overline{A'})/2, \qquad f = c(A+A^* - (A'+\overline{A'}))/2\delta$$

$$m = c((A-A^*)\sqrt{\gamma} + (A'-\overline{A'})\sqrt{\gamma'})/2,\ n = c((A-A^*)\sqrt{\gamma} - (A'-\overline{A'})\sqrt{\gamma'})/2\delta$$

PROOF

The result follows easily from definition 3.3 and (16). //

The calculation of U_1 required in step 4 of algorithm 3.5 is based on theorem 3.21 plus the following result.

THEOREM 3.22

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$, $\gamma > 0$, be a type IIb quartic field with U_1, δ_1, U_f as in theorem 3.20. Then $U_1^* = U_1^{-1}$ and

$$\text{either } U_1 = |U_f|, \quad U_1' = s\sqrt{U_f N_{\delta_1}(U_f) U_f^{-1}}, \quad s \in \{-1,1\}$$

$$\text{or } U_1 = |U_f|^2, \quad U_1' = U_f \overline{N_{\delta_1}(U_f) U_f^{-1}}$$

PROOF

The expressions for U_1^*, U_1 follow directly from theorems 3.19, 3.20. To prove the result for U_1' we first note that for $U = U_f U_f'$ we have $U' = U_f U_f'^*$. This result is proved using the same approach as in the proof of theorem 3.20. Now $U_f'^* = \overline{U_f^*}$, $U_f U_f^* = N_{\delta_1}(U_f)$ and so

$$U' = U_f U_f^{*\prime} = U_f \overline{N_{\delta_1}(U_f) U_f^{-1}}$$

The possibilities for U_1' now follow easily since $(U_1')^2 = U'$ if $U_1^2 = |U_f|^2 = U$ and $U_1' = U'$ if $U_1 = |U_f|^2 = U$. //

Thus to find U_1 we first test

$$A = |U_f|, \quad A^* = |U_f|^{-1}, \quad A' = s\sqrt{U_f \overline{N_{\delta_1}(U_f) U_f^{-1}}}$$

with $s = 1$ to see if e, f, m, n in theorem 3.21 are rational integers. If this test is successful then $U_1 = |U_f|$ and e, f, m, n give the coefficients of U_1. If the test fails then we try $s = -1$. Finally if both $s = 1$, $s = -1$ fail to produce U_1 then we can conclude that

$U_1 = |U_f|^2$ and the coefficients of U_1 are obtained by taking

$$A = |U_f|^2,\ A^* = |U_f|^{-2},\ A' = U_f\overline{N_{\delta_1}(U_f)U_f^{-1}}$$

in theorem 3.21. A more precise statement of these ideas is given in the following algorithm.

ALGORITHM 3.5A

The calculation of U_1 required in step 4 of algorithm 3.5 can be carried out as follows.

1 Set $x = |U_f|$, $y = \sqrt{U_f\overline{N_{\delta_1}(U_f)U_f^{-1}}}$, $s = 1$

2 If $e = c(x+x^{-1}+s(y+\bar{y}))/2 \notin Z$ go to 7

3 If $f = c(x+x^{-1}-s(y+\bar{y}))/2\delta \notin Z$ go to 7

4 If $m = c((x-x^{-1})\sqrt{\gamma}+s(y-\bar{y})\sqrt{\gamma'})/2 \notin Z$ go to 7

5 If $n = c((x-x^{-1})\sqrt{\gamma}-s(y-\bar{y})\sqrt{\gamma'})/2\delta \notin Z$ go to 7

6 We have $U_1 = (\alpha+\beta/\sqrt{\gamma})/2$, $\alpha = (e+f\delta)/c$, $\beta = (m+n\delta)/c$ and we stop

7 If $s = 1$ set $s = -1$ and go to 2

8 Set $x = |U_f|^2$, $y = U_f\overline{N_{\delta_1}(U_f)U_f^{-1}}$, $s = 1$, go to 2 //

The testing of whether or not $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ required in step 5 of algorithm 3.5 is generally much simpler than the calculation of U_1.

THEOREM 3.23

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$, $\gamma > 0$, be a type IIb quartic field. If $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ then $N(\varepsilon(d)) = -1$, $N(\gamma) = -k^2$ with $k \in Z$, and

$$\sqrt{\varepsilon(d)} = \beta/\sqrt{\gamma},\ \beta \in Z(\delta)$$

$$(\sqrt{\varepsilon(d)})^* = -\sqrt{\varepsilon(d)}$$

$$(\sqrt{\varepsilon(d)})' = si/\sqrt{\varepsilon(d)},\ s \in \{-1,1\}$$

PROOF

Suppose $\sqrt{\varepsilon(d)} \in Q(\Gamma)$. Then $Q(\Gamma) = Q(\delta)(\sqrt{\varepsilon(d)})$. Now as was noted following theorem 3.1 this implies $\gamma\varepsilon(d) = \beta^2$, $\beta \in Z(\delta)$. Since $N(\gamma) < 0$ the conditions which $N(\varepsilon(d))$, $N(\gamma)$ must satisfy are clear. Since $\beta^2 = (-\beta)^2$ we can choose β so that $\sqrt{\varepsilon(d)} = \beta/\sqrt{\gamma}$ and the form of $(\sqrt{\varepsilon(d)})^*$ follows trivially. Note that

$$\varepsilon(d)' = ((\sqrt{\varepsilon(d)})^2)' = ((\sqrt{\varepsilon(d)})')^2$$

and so $(\sqrt{\varepsilon(d)})' = s\sqrt{\varepsilon(d)'}$ with $s \in \{-1,1\}$. Since $N(\varepsilon(d)) = -1$ we have $\sqrt{\varepsilon(d)'} = \sqrt{N(\varepsilon(d))/\varepsilon(d)} = i/\sqrt{\varepsilon(d)}$. //

Thus to check if $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ we first check $N(\varepsilon(d))$, $N(\gamma)$. If these norms satisfy the required conditions then we test

$$A = \sqrt{\varepsilon(d)},\ A^* = -A,\ A' = -\overline{A}' = si/\sqrt{\varepsilon(d)}$$

to see if m,n in theorem 3.21 are rational integers. Note that we can ignore e,f in theorem 3.21 since we will find e = f = 0 regardless of whether or not $\sqrt{\varepsilon(d)} \in Q(\Gamma)$.

ALGORITHM 3.5B

The testing of whether or not $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ required in step 5 of algorithm 3.5 can be carried out as follows.

1 If $N(\varepsilon(d)) = 1$ or $N(\gamma) \neq -k^2$, $k \in Z$ then $\sqrt{\varepsilon(d)} \in Q(\Gamma)$ and we stop

2 Set $x = \sqrt{\varepsilon(d)}$, $y = i/x$, $s = 1$

3 If $m = c(x\sqrt{\gamma} + sy\sqrt{\gamma'}) \notin Z$ go to 6

4 If $n = c(x\sqrt{\gamma} - sy\sqrt{\gamma'})/\delta \notin Z$ go to 6

5 We have $\sqrt{\varepsilon(d)} = \beta/\sqrt{\gamma} \in Q(\Gamma)$, $\beta = ((m/2) + (n/2)\delta)/c$ and we stop

6 If $s = 1$ then set $s = -1$ and go to 3

7 We have $\sqrt{\varepsilon(d)} \notin Q(\Gamma)$ and we stop //

Note that since theorem 3.21 uses the form $A = (\alpha+\beta/\sqrt{\gamma})/2$ the coefficients in steps 3, 4 will be those of 2β when $\varepsilon(d) = \beta/\sqrt{\gamma} \in Q(\Gamma)$. This is the reason for the division by 2 in step 5.

The testing of whether or not $\sqrt{\varepsilon(d)U_1} \in Q(\Gamma)$ required in step 6 of algorithm 3.5 is similar to the previous two cases.

THEOREM 3.24

Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$, $\gamma > 0$, be a type IIb quartic field with U_1 as in theorem 3.19. If $\sqrt{\varepsilon(d)U_1} \in Q(\Gamma)$ then

$$(\sqrt{\varepsilon(d)U_1})^* = N(\varepsilon(d))\varepsilon(d)/\sqrt{\varepsilon(d)U_1}$$

$$(\sqrt{\varepsilon(d)U_1})' = s\sqrt{\varepsilon(d)'U_1'}, \quad s \in \{-1,1\}$$

PROOF

Suppose $U = \sqrt{\varepsilon(d)U_1} \in Q(\Gamma)$. Nagell [1961, p363] gives the first result and the second result follows from the relationship

$$(U')^2 = (U^2)' = \varepsilon(d)'U_1' \qquad //$$

ALGORITHM 3.5C

The testing of whether or not $\sqrt{\varepsilon(d)U_1} \in Q(\Gamma)$ required in step 6 of algorithm 3.5 can be carried out as follows.

1 Set $x = \sqrt{\varepsilon(d)U_1}$, $y = N(\varepsilon(d))\varepsilon(d)/x$, $z = \sqrt{\varepsilon(d)'U_1'}$, $s = 1$

2 If $e = c(x+y+s(z+\bar{z}))/2 \notin Z$ go to 7

3 If $f = c(x+y-s(z+\bar{z}))/2\delta \notin Z$ go to 7

4 If $m = c((x-y)\sqrt{\gamma}+s(z-\bar{z})\sqrt{\gamma'})/2 \notin Z$ go to 7

5 If $n = c((x-y)\sqrt{\gamma}-s(z-\bar{z})\sqrt{\gamma'})/2\delta \notin Z$ go to 7

6 We have $\sqrt{\varepsilon(d)U_1} = (\alpha+\beta/\sqrt{\gamma})/2 \in Q(\Gamma)$, $\alpha = (e+f\delta)/c$, $\beta = (m+n\delta)/c$ and we stop

7 If $s = 1$ then set $s = -1$ and go to 2

8 We have $\sqrt{\varepsilon(d)U_1} \notin Q(\Gamma)$ and we stop //

This completes the testing procedures required by algorithm 3.5. Note that in practice we must adopt the same approach as in algorithms 3.3, 3.4 as far as the calculation of e, f, m, n is concerned. That is sufficient precision is used so as to ensure that these quantities are correct to at least 3D and initial decisions are based on whether or not the results are rational integers to within rounding error. As a final check we add the following norm tests.

Algorithm 3.5A If $x = |U_f|$ when step 6 is reached then check that we have

$$U_1 U_1^* = (\alpha^2 - \beta^2/\gamma)/4 = 1$$

Note that this check is not required when $x = |U_f|^2$ since $|U_f|^2$ is always a unit of $Q(\Gamma)$.

Algorithm 3.5B If step 5 is reached then check that we have

$$\beta^2 = \gamma\varepsilon(d)$$

Algorithm 3.5C If step 6 is reached check that we have

$$\sqrt{\varepsilon(d)U_1}(\sqrt{\varepsilon(d)U_1})^* = (\alpha^2 - \beta^2/\gamma)/4 = N(\varepsilon(d))\varepsilon(d)$$

Note that the calculations required by algorithm 3.5 involve $O(\ell n(|U_f|\varepsilon(d)))$ precision arithmetic and so multiprecision arithmetic is required when $|U_f|$ or $\varepsilon(d)$ is large.

The results of this section are now illustrated in the following example.

EXAMPLE 3.6

(a) Let $Q(\Gamma) = Q(\sqrt[4]{82})$. Then $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ where $\delta = \sqrt{82}$,

$\gamma = \delta$. We calculate a fundamental system of units for $Q(\Gamma)$ using algorithm 3.5.

The simple continued fraction algorithm gives $\varepsilon(82) = 9 + \delta$. We have $a = 0$, $c = 1$, $N(\gamma) = -82$ and so $\kappa = 2\delta_1$, $\delta_1 = \sqrt{-82}$. Thus $Q(\delta_1)(\sqrt{\kappa}) = Q(\sqrt[4]{-328})$. Using (15) we determine $\underline{Z}(\delta_1)(\sqrt{\kappa})$ as being of form 4. We therefore use algorithm 3.4 to calculate

$$U_f = ((18+2\delta_1) + (12\delta_1)/\sqrt{\kappa})/2, \quad N_{\delta_1}(U_f) = -1$$

Algorithm 3.5A is now used to calculate U_1 for $Q(\Gamma)$. The possibility $U_1 = |U_f|$ is rejected since

$$e \approx 13.495 \notin Z \quad \text{for } s = 1$$

$$e \approx 12.079 \notin Z \quad \text{for } s = -1$$

in step 2 of algorithm 3.5A. Consequently $|U_1| = |U_f|^2$ and steps 2, 3, 4, 5 give the coefficients of U_1. We have

$$U_1 = ((326+36\delta) + (984+108\delta)/\sqrt{\gamma})/2$$

Since $N(\gamma) \neq -k^2$, $k \in Z$ algorithm 3.5B concludes $\sqrt{\varepsilon(82)} \notin Q(\Gamma)$. We therefore apply algorithm 3.5C. For $s = 1$ we find $e \approx 54.332 \notin Z$. However for $s = -1$ we find

$$e \approx 54.000, \; f \approx 6.000, \; m \approx 164.000, \; n \approx 18.000$$

A norm check confirms that this is indeed a unit of $Q(\Gamma)$ and so we have

$$\sqrt{\varepsilon(82)U_1} = ((54+6\delta) + (164+18\delta)/\sqrt{\gamma})/2 \in Q(\Gamma)$$

Thus $\{\varepsilon(82), \sqrt{\varepsilon(82)U_1}\}$ is a fundamental system of units of $Q(\Gamma)$. (This system is equivalent to the system $\{U_1, \sqrt{\varepsilon(82)^{-1}U_1}\}$ given by Ljunggren [1934, p15]).

(b) Let $Q(\Gamma) = Q(\delta)(\sqrt{\gamma})$ with $\delta = \sqrt{221}$, $\gamma = -6 + 4\delta$. We have $\varepsilon(221) = (15+\delta)/2$. Now $N(\gamma) = -3500 = -35.10^2$ and so we set

$$\kappa = -3 + 5\delta_1, \ \delta_1 = \sqrt{-35} \quad \text{(note that } \ell = 2\text{)}$$

In example 3.2b we saw that for $Q(\delta_1)(\sqrt{\kappa})$ we can take

$$U_f = (62-25\delta_1) + ((-25-9\delta_1)/2)\sqrt{\kappa}$$

Algorithm 3.5A shows that

$$U_1 = |U_f|^2 = ((51438+3460\delta) + (376100+25300\delta)/\sqrt{\gamma})/2$$

Since $N(\varepsilon(221)) = 1$ algorithm 3.5B concludes $\sqrt{\varepsilon(221)} \notin Q(\Gamma)$. Finally algorithm 3.5C shows that $\sqrt{\varepsilon(221)U_1} \notin Q(\Gamma)$. Thus $\{\varepsilon(221), U_1\}$ is a fundamental system of units for $Q(\Gamma)$. //

We finish this section by briefly noting several relevant results from the literature. Ljunggren [1934] gives a table which contains the unit U_1 for many of the type IIb quartic fields $Q(\sqrt[4]{D})$ with $D \in Z^+$, $1 < D < 100$. This unit plus either $\varepsilon(d)$ or $\sqrt{\varepsilon(d)^{-1}U_1}$ (d being the square-free kernal of D) form a fundamental system of units of $Q(\sqrt[4]{D})$. However it is not clear how Ljunggren has calculated many of these units. Shanks [1977] gives several corrections to Ljunggren's table. (See also the comments in the final paragraphs of the previous section). It is not clear if Shanks' method is applicable to more general type IIb quartic fields.

The Jacobi-Perron Algorithm (see Bernstein [1971]) has limited success in calculating units of type IIb quartic fields. Works such as Bernstein [1977] (summarising the work of various authors) and Frei and Levesque [1979] list explicit units for a number of infinite classes of algebraic fields which include some type IIb quartic fields of the form $Q(\sqrt[4]{D}), D \in Z^+$. (In some cases these units are obtained from the Jacobi-Perron Algorithm). However this type of result only deals with a very small subset of the set of type IIb quartic fields.

We also note that general algorithms such as those of Billevič (see Steiner and Rudman [1976]) and Berwick (see Rudman and Steiner [1978]) can be used to calculate units of type IIb quartic fields.

To conclude we note that it is quite probable that a direct approach to calculating fundamental systems of units of type IIb quartic fields (rather than via type I quartic fields) will be more efficient. However we have presented the ideas in this section as an interesting side issue (and bonus) to the main thrust of this thesis, that is the calculation of units of type I quartic fields.

ADDENDUM

It has been pointed out by the overseas examiner that

BUCHMANN, J. [1982]

Zahlengeometrische Kettenbruchalgorithmen Zur Einheitberechnung. Inaugural Dissertation, Zur Erlangung den Doktorgrades der Mathematisch Naturwissenschaflichen Fakultät der Universität zu Köln, Köln

covers work of a similar nature to that covered in this thesis. When this work becomes available I intend to make a comparison of the two works.

Neville Jeans

REFERENCES

ADAMS, W.W. and GOLDSTEIN, L.J. [1976]

Introduction to Number Theory. Prentice-Hall, New Jersey.

AMARA, H. [1981]

Groupe Des Classes Et Unite Fondamentale Des Extensions Quadratiques Relatives A Un Corps Quadratique Imaginaire Principal. *Pacific Journal of Mathematics*, Vol 96, No 1, pp1-12.

ARWIN, A. [1926]

Einige periodische Kettenbruchentwicklungen. *Journal für die reine und angewandte Mathematik*, Band 155, pp111-128.

BERNSTEIN, L. [1971]

Lecture Notes in Mathematics, Vol 207, The Jacobi-Perron Algorithm Its Theory and Application. Springer-Verlag, New York.

BERNSTEIN, L. [1977]

Gaining Units From Units. *Canadian Journal of Mathematics*, No 1, pp93-106.

BERWICK, W.E.H. [1913]

The Classification of Ideal Numbers that Depend on a Cubic Irrationality. *Proceedings of the London Mathematical Society* (2) 12 (1913), pp393-429.

BILLEVIČ, K.K. [1956]

On units of algebraic fields of third and fourth degrees. *Matematičeskiĭ Sbornik*, Vol 40 (82), pp123-136 (Russian).

BOREVICH, Z.I. and SHAFAREVICH I.R. [1966]

Number Theory. Academic Press, New York.

BRENTJES, A.J. [1981]

Multi-dimensional Continued Fraction Algorithms. Mathematisch Centrum, Amsterdam.

BUELL, D.A., WILLIAMS, H.C. and WILLIAMS, K.S. [1977]

On the Imaginary Bicyclic Biquadratic Fields with Class-Number 2. *Mathematics of Computation*, Vol 31, No 140, pp1034-1042.

CHRYSTAL, G. [1959]

Algebra: An Elementary Text-book for the Higher Classes of Secondary Schools and for Colleges, Part II. 6th ed. Chelsea Publishing Company, New York.

COHN, H. [1962]

A Second Course in Number Theory. John Wiley and Sons, New York.

FEJES TÓTH L. [1953]

Lagerungen in der Ebene, auf der Kugel und im Raum. (Grundlehren der mathematischen Wissenschaften, 65). Springer-Verlag, Berlin.

FJELLSTEDT, L. [1953]

On a class of Diophantine equations of the second degree in imaginary quadratic fields. *Arkiv för Matematik,* Band 2, nr 24 pp435-461.

FREI, G. [1982]

Fundamental Systems of Units in Biquadratic Parametric Number Fields. *Journal of Number Theory,* Vol 15, pp295-303.

FREI, G. and LEVESQUE, C. [1979]

Independent systems of units in certain algebraic number fields. *Journal für die reine und angewandte Mathematik,* Band 311, pp116-144.

HARDY, G.H. and WRIGHT, E.M. [1979]

An Introduction to the Theory of Numbers. 5th ed, Oxford University Press, Oxford.

HENDY, M.D. and JEANS, N.S. [1981]

The Jacobi-Perron Algorithm in Integer Form. *Mathematics of Computation,* Vol 36, No 154, pp565-574.

HILBERT, D. [1894]

Über den Dirichletschen biquadratischen Zahlkörper. *Mathematische Annalen,* 45, pp309-340.

HURWITZ, A. [1887]

Über die Entwicklung complexer Grössen in Kettenbrüche. *Acta Mathematica,* Vol 11, pp187-200.

HURWITZ, J. [1902]

Über die Reduction per binären quadratischen Formen mit complexen Coefficienten und Variablen. *Acta Mathematica*, Vol 25, pp231-290.

JEANS, N.S. [1978]

Units in Some Algebraic Number Fields. Thesis, M.Sc., Massey University.

JEANS, N.S. and HENDY, M.D. [1978]

Determining the Fundamental Unit of a Pure Cubic Field given any Unit. *Mathematics of Computation*, Vol 32, No 143, pp925-935.

KUBOTA, T. [1953]

Über die Beziehung der Klassenzahlen der Unterkörper des bizyklischen biquadratischen Zahlkörpers. *Nagoya Mathematical Journal*, Vol 6, pp119-127.

KUBOTA, T. [1956]

Über den bizyklischen biquadratischen Zahlkörper. *Nagoya Mathematical Journal*, Vol 10, pp65-85.

KURODA, S. [1943]

Über den Dirichletschen Körper. *Journal of the Faculty of Science, Imperial University of Tokyo. Section I A Mathematics*, Vol 4, pp383-406.

LAKEIN, R.B. [1971]

Class Numbers and Units of Complex Quadratic Fields. pp167-172 in *Computers in Number Theory*. Atkin, A.O.L. and Birch, B.J. (editors), Academic Press, London.

LAKEIN, R.B. [1974]

Computation of the Ideal Class Group of Certain Complex Quartic Fields. *Mathematics of Computation*, Vol 28, No 127, pp839-846.

LAKEIN, R.B. [1975]

Computation of the Ideal Class Group of Certain Complex Quartic Fields II. *Mathematics of Computation*, Vol 29, No 129, pp137-144.

LeVEQUE, W.J. [1977]

Fundamentals of Number Theory. Addison-Wesley, Reading, Massachussetts.

LEVESQUE, C. [1981]

Systemes fondamentaux d'unites de certains composes de deux corps quadratiques, I. *Canadian Journal of Mathematics,* Vol 33, No 4, pp937-945.

LJUNGGREN, W. [1934]

Über die Lösung einiger unbestimmten Gleichungen vierten Grades. *Avhandlinger (Norske Videnskapsakademi I Oslo) Matermatisk Naturvidenskabelige Klasse,* No 14.

MACKENZIE, R. and SCHEUNEMAN, J. [1971]

A Number Field without a Relative Integral Basis. *American Mathematical Monthly,* 78, pp882-883.

NAGELL, T. [1961]

Sur quelques questions dans la théorie des corps biquadratiques. *Arkiv för Matematik,* Band 4, nr 26, pp347-376.

PARRY, C.J. [1980]

A genus theory for quartic fields. *Journal für die reine und angewandte Mathematik,* Band 314, pp40-71.

RICHMAN, F. [1971]

Number Theory: An Introduction to Algebra. Brooks/Cole, Belmont, California.

RUDMAN, R.J. and STEINER, R.P. [1978]

A Generalization of Berwick's Unit Algorithm. *Journal of Number Theory,* 10, pp16-34.

SHANKS, D. [1977]

Table Errata for W. Ljunggren. *Mathematics of Computation,* Vol 31, No 140, pp1049-1050.

STEIN, A. [1927]

Die Gewinnung der Einheiten in gewissen relativ-quadratischen Zahlkörpern durch das J. Hurwitzsche Kettenbruchverfahran. *Journal für die reine und angewandte Mathematik*, Band 156, pp69-92.

STEINER, R. and RUDMAN, R. [1976]

On an Algorithm of Billevich for finding Units in Algebraic Number Fields. *Mathematics of Computation*, Vol 30, pp598-609.

STEWART, I.N. and TALL, D.O. [1979]

Algebraic Number Theory. Chapman and Hall, London.

SZEKERES, G. [1970]

Muldidimensional Continued Fractions. *Annales Universitatis Scientiarum Budapestinensis De Rolando Eötvös Nominatae Sectio Mathematics*, Vol 13, pp113-140.

VORONOI, G.F. [1896]

On a Generalization of the Algorithm of Continued Fractions. Doctoral Dissertation, Warsaw (Russian).

WILLIAMS, H.C. [1980]

Improving the Speed of Calculating the Regulator of Certain Pure Cubic Fields. *Mathematics of Computation*, Vol 35, No 152, pp1423-1434.

WILLIAMS, H.C. and BUHR, P.A. [1979]

Calculation of the regulator of $Q(\sqrt{d})$ by use of the nearest integer continued fraction algorithm. *Mathematics of Computation*, Vol 33, pp369-381.

WILLIAMS, H.C., CORMACK, G. and SEAH, E. [1980]

Calculation of the Regulator of a Pure Cubic Field. *Mathematics of Computation*, Vol 34, No 150, pp567-611.

WILLIAMS, K.S. [1970]

Integers of Biquadratic Fields. *Canadian Mathematical Bulletin*, 13 (4), pp519-526.