On the Rank of Universal Quadratic Forms
over Real Quadratic Fields

Valentin Blomer and Vítězslav Kala[1]

Abstract. We study the minimal number of variables required by a totally positive definite diagonal universal quadratic form over a real quadratic field $\mathbb{Q}(\sqrt{D})$ and obtain lower and upper bounds for it in terms of certain sums of coefficients of the associated continued fraction. We also estimate such sums in terms of D and establish a link between continued fraction expansions and special values of L-functions in the spirit of Kronecker's limit formula.

2010 Mathematics Subject Classification: Primary 11E12; Secondary 11R11, 11A55.
Keywords and Phrases: universal quadratic form, real quadratic number field, continued fraction, additively indecomposable integer

1   Introduction

The story of universal quadratic forms began with the four square theorem that was proved by Lagrange and states that the quadratic form $x^2 + y^2 + z^2 + w^2$ represents all positive integers. This was followed by a large number of further results over the ring of integers $\mathbb{Z}$, such as the classification of all universal quaternary diagonal forms by Ramanujan and Dickson, the 15-theorem of Conway, Miller and Schneeberger (see [Bh] for a beautiful proof) and the 290-theorem of Bhargava and Hanke [BH].
Maaß [Ma] proved that the sum of three squares is universal over $\mathbb{Q}(\sqrt{5})$, and then Siegel [Si] showed that over every other number field, the sum of any

number of squares is not universal. This turned the attention to other quadratic forms: Chan, Kim, and Raghavan [CKR] studied ternary universal forms over real quadratic fields, and for example showed that diagonal ternary universal forms exist only over $\mathbb{Q}(\sqrt{2})$, $\mathbb{Q}(\sqrt{3})$, and $\mathbb{Q}(\sqrt{5})$. There were numerous other results over specific real quadratic fields, e.g., Sasaki [Sa] found all universal quaternary forms over $\mathbb{Q}(\sqrt{13})$. For more general discriminants, Kim [Ki1, Ki2] proved that there are only finitely many real quadratic fields that have a diagonal universal form in 7 variables, but also constructed an 8-ary diagonal universal form over each field $\mathbb{Q}(\sqrt{n^2-1})$ (when n^2-1 is squarefree).

For a general number of variables m, the present authors [BK, Ka1] constructed infinitely many real quadratic fields that do not admit m-ary universal quadratic forms. Using very different techniques, Yatsyna [Ya] recently extended these results to the case of number fields that possess units of every signature.

However, still not much is known about the number of variables required by a general (diagonal) universal form over a real quadratic field: this is the goal of the present paper. We will focus on totally positive definite diagonal quadratic forms of arity (or rank) m over real quadratic fields $K = \mathbb{Q}(\sqrt{D})$, i.e., forms $Q(x_1, \ldots, x_m) = a_1 x_1^2 + \cdots + a_m x_m^2$, where $a_1, \ldots, a_m$ are totally positive elements of the ring of integers $\mathcal{O}_K$. Such a form is universal if it represents all totally positive integers α, i.e., if $\alpha = Q(x_1, \ldots, x_m)$ for some $x_1, \ldots, x_m \in \mathcal{O}_K$. Let us denote by $m_{\mathrm{diag}}(D)$ the smallest integer m such that there is an m-ary diagonal universal form over K. We introduce the following notation. For squarefree $D > 1$ let

$$\omega_D = \begin{cases} \sqrt{D} & \text{if } D \equiv 2, 3 \\ \frac{1+\sqrt{D}}{2} & \text{if } D \equiv 1 \end{cases} \pmod{4}, \tag{1.1}$$

$\Delta \in \{D, 4D\}$ the discriminant of $\mathcal{O}_K$ and $\omega_D = [u_0, \overline{u_1, u_2, \ldots, u_{s-1}, u_s}]$ the periodic continued fraction expression with s minimal. We define

$$M_D = \begin{cases} u_1 + u_3 + \cdots + u_{s-1} & \text{if } s \text{ is even,} \\ 2u_0 + u_1 + u_2 + \cdots + u_{s-1} & \text{if } s \text{ is odd and } D \equiv 2, 3 \pmod 4, \\ 2u_0 + u_1 + u_2 + \cdots + u_{s-1} - 1 & \text{if } s \text{ is odd and } D \equiv 1 \pmod 4. \end{cases} \tag{1.2}$$

If s is odd, equivalently if $\mathcal{O}_K$ has a unit with negative norm, then

$$M_D = u_1 + \ldots + u_s. \tag{1.3}$$

For $\varepsilon > 0$ define $M_{D,\varepsilon}^*$ as the sum in (1.2), but ranging only over coefficients $u_i \geq D^{1/8+\varepsilon}$.

Our first result is then the following:

THEOREM 1. *With the above notation we have*

$$\max\left(\frac{M_D}{\kappa s}, C_\varepsilon M_{D,\varepsilon}^*\right) \leq m_{\mathrm{diag}}(D) \leq 8M_D$$

for any $\varepsilon > 0$, where $C_\varepsilon > 0$ is a constant (depending only on ε) and $\kappa = 2$ if s is odd and $\kappa = 1$ otherwise.

We will prove this result as Theorems 10, 12, and 14. All of these proofs are based on studying the representability of additively indecomposable integers of $\mathbb{Q}(\sqrt{D})$, which can be nicely characterized in terms of convergents and semi-convergents to the corresponding continued fraction. To show the upper bound, we generalize Kim's result [Ki2] and construct an explicit diagonal universal form, whose coefficients are (some of) these indecomposable integers. The lower bound hinges on showing that sufficiently many indecomposables (essentially) have to appear as the coefficients of any diagonal universal form. Comparing our two lower bounds $\frac{M_D}{\kappa s}$ and $C_\varepsilon M_{D,\varepsilon}^*$, the first one is clearly larger for small values of D or s. However, one can estimate the size of s (cf. (6.4)) to see that the second bound is larger for most D's.

Of course one would like to know how M_D and $M_{D,\varepsilon}^*$ behave asymptotically with respect to D or equivalently Δ. In this respect, we will prove in Section 6 the following.

THEOREM 2. *We have $M_D \leq c\sqrt{\Delta}(\log \Delta)^2$ for an absolute constant $c > 0$. If s is odd, equivalently if $\mathcal{O}_K$ has a unit of negative norm, we have $M_{D,\varepsilon}^* \geq \sqrt{\Delta}$ for every $\varepsilon < 1/8$.*

Under the Generalized Riemann Hypothesis, the upper bound can be sharpened somewhat (cf. (6.3)). The lower bound for odd s follows simply from the fact that $u_0 = \lfloor \omega_D \rfloor \geq \frac{1}{2}\sqrt{\Delta}$.

With more analytic work, one can even establish an asymptotic formula for the sum of convergents, which may be of independent interest. It can be viewed as a variation of Kronecker's limit formula for real quadratic fields and highlights the fascinating connection between special L-values and continued fractions. The following result is proved in Corollary 18, which in view of (1.3) gives an asymptotic formula for M_D if s is odd.

THEOREM 3. *As $D \to \infty$, we have*

$$\sum_{i=1}^{s} u_i \sim \frac{\sqrt{\Delta}}{\zeta^{(\Delta)}(2)}\Big(L(D) + \frac{1}{h}L(1, \chi_\Delta)\log\sqrt{D}\Big),$$

where $h = h_D$ is the class number, $\zeta^{(\Delta)}(s)$ is the Riemann zeta function with Euler factors at primes dividing Δ removed, χ_Δ is the usual quadratic character associated with the fundamental discriminant Δ, and $L(D)$ (defined in (5.2)) is the constant Taylor coefficient of the ζ-function associated to the class of principal ideals.

Kronecker's limit formula is concerned with finding a closed expression for $L(D)$ (and more general functions). Zagier and Hirzebruch observed that for real quadratic fields there is a connection between $L(D)$ and the coefficients of the continued fraction of ω_D. The exact formula in [Za1, Corollary 2] (which

is derived by a completely different method than Theorem 3), however, seems to be hard to use to obtain any sort of asymptotic statement. The beautiful formula [Za2, Satz 2, §14], on the other hand, is of different nature, since it treats the alternating sum $\sum_{i=1}^{s}(-1)^i u_i$, cf. [Za2, p. 131] (and gives in particular no information if s is odd). Yet another variation of the connection between special values of class group L-functions for real quadratic fields and continued fractions can be found in a nice paper of Biró and Granville [BG, Theorem 1].

Most of what we do in this paper can probably be generalized to non-maximal orders in $\mathcal{O}_K$, but in order to avoid technical subtleties in particular in the analytic argument, we concentrate on the most natural case of the maximal order.

2   Preparation: Indecomposables

2.1   Basic notation

Throughout the paper, we shall use the following notation: Let $K = \mathbb{Q}(\sqrt{D})$ with squarefree $D > 1$ and discriminant Δ. For $\alpha = x + y\sqrt{D} \in \mathbb{Q}(\sqrt{D})$ we denote its conjugate as $\alpha' = x - y\sqrt{D} \in \mathbb{Q}(\sqrt{D})$ and its norm as $N(\alpha) = x^2 - Dy^2$. An element α is totally positive if $\alpha > 0$ and $\alpha' > 0$; if $\alpha - \beta$ is totally positive, we write $\alpha \succ \beta$. We denote by $\mathcal{O}_K$ the ring of integers of K and by $\mathcal{O}_K^+$ the semiring of totally positive integers. Let ω_D be as in (1.1), so that

$$-\omega_D' = \begin{cases} \sqrt{D} & \text{if } D \equiv 2,3 \\ \frac{-1+\sqrt{D}}{2} & \text{if } D \equiv 1 \end{cases} \pmod{4},$$

and $\mathcal{O}_K = \mathbb{Z}[\omega_D]$. We recall from the introduction that $\omega_D = [u_0, \overline{u_1, u_2, \ldots, u_{s-1}, u_s}]$. Note that $u_s = 2u_0$ when $D \equiv 2,3 \pmod 4$ and $u_s = 2u_0 - 1$ when $D \equiv 1 \pmod 4$ (which we used to go from (1.2) to (1.3)). We also know that the sequence $(u_1, u_2, \ldots, u_{s-1})$ is symmetric, i.e., $u_i = u_{s-i}$.

Let $\frac{p_i}{q_i} := [u_0, \ldots, u_i]$ be the ith convergent to ω_D (where p_i, q_i are coprime positive integers), and $\alpha_i = p_i - q_i\omega_D'$ the corresponding element of $\mathcal{O}_K$, which by small abuse of notation we also call a convergent. We have $p_{i+1} = u_{i+1}p_i + p_{i-1}$, $q_{i+1} = u_{i+1}q_i + q_{i-1}$, and $\alpha_{i+1} = \alpha_i q_i + \alpha_{i-1}$ (with initial conditions $p_{-1} := 1$, $p_0 = k$, $q_{-1} := 0$, $q_0 = 1$). Note that $\alpha_{-1} = 1$, and that $\alpha_i \succ 0$ if and only if i is odd.

Let ϵ be the *totally positive* fundamental unit, $N(\epsilon) = 1$. Then we have $\epsilon = \alpha_{s-1}$ or $\epsilon = \alpha_{2s-1}$ when s is even or odd, respectively.

By a semi-convergent to ω_D we mean a fraction of the form $\frac{p_i + r p_{i+1}}{q_i + r q_{i+1}}$ with $0 \leq r \leq u_{i+2}$ ($i \geq -1$). Note that each convergent is also a semi-convergent and that if we take $r = u_{i+2}$, the fraction is just p_{i+2}/q_{i+2}. We denote the corresponding element of $\mathcal{O}_K$ (which we also call a semi-convergent) by $\alpha_{i,r} := \alpha_i + r \alpha_{i+1}$ (again for $i \geq -1$ and $0 \leq r \leq u_{i+2}$). A semi-convergent $\alpha_{i,r}$ is a convex combination of α_i and α_{i+2}. These are either both totally positive, or none of them is, and hence $\alpha_{i,r}$ is totally positive if and only if i is odd.

Denote by S the set of all totally positive semi-convergents and their conjugates, i.e.,

$$ S := \{ \alpha_{i,r}, \alpha'_{i,r} \mid i \geq -1 \text{ odd}, 0 \leq r < u_{i+2} \}, $$

and by S_0 its subset of elements $\sigma \in S$ satisfying $\epsilon > \sigma \geq \sigma' > 0$ (note that in the definition of S, all the elements are distinct, except for $\alpha_{-1,0} = 1 = \alpha'_{-1,0}$). The set S is closed under multiplication by ϵ^k for any $k \in \mathbb{Z}$, and each element of S can be uniquely written in the form $\sigma_0 \epsilon^k$ for some $\sigma_0 \in S_0$ and $k \in \mathbb{Z}$. The cardinality of S_0 is clearly $u_1 + u_3 + \cdots + u_{s-1}$ if s is even and $u_1 + u_3 + \cdots + u_s + u_{s+2} + \cdots + u_{2s-1}$ if s is odd, because S_0 consists precisely of all the elements $\alpha_{i,r}$ with odd $-1 \leq i \leq s - 3$ ($2s - 3$, resp.). Note that this equals M_D as defined in (1.2), as $u_s = 2u_0$ or $u_s = 2u_0 - 1$ when $D \equiv 2, 3 \pmod 4$ or $1 \pmod 4$ and $u_{s+i} = u_i$ for $i \geq 1$.

A totally positive integer $\alpha \in \mathcal{O}_K^+$ is (additively) indecomposable if it cannot be decomposed as the sum of two totally positive elements, i.e., if $\alpha \neq \beta + \gamma$ for $\beta, \gamma \in \mathcal{O}_K^+$. By a classical theorem (see, e.g., [Pe, §16] or [DS, Theorem 2]), semi-convergents (and their conjugates) are exactly all indecomposable elements, i.e., S is the set of all indecomposable integers.

In later sections we will use the following asymptotic notation: For real functions $f(x), g(x)$, we write $f \sim g$ if $\lim_{x \to \infty} f(x)/g(x) = 1$, $f \ll g$ (or $g \gg f$) if there are $c > 0$ and x_0 such that $|f(x)| < cg(x)$ for all $x > x_0$, and $f \asymp g$ if $f \ll g$ and $f \gg g$.

2.2 Estimates

We will need certain estimates on the sizes of indecomposables and their norms. These are mostly classical; in the case $D \equiv 2, 3 \pmod 4$ they have appeared for example in [Ka2]. We will need them also when $D \equiv 1 \pmod 4$, and so we collect the required results here, giving proofs only for this case.

We will first need to introduce some additional notation. For a convergent α_i, we set

$$ N_i := |N(\alpha_i)| = (-1)^{i+1} N(\alpha_i) = \begin{cases} |p_i^2 - p_i q_i - q_i^2 \frac{D-1}{4}|, & D \equiv 1 \,(\mathrm{mod}\ 4), \\ |p_i^2 - D q_i^2|, & D \equiv 2, 3 \,(\mathrm{mod}\ 4). \end{cases} $$

Recall that we have $p_{i+1}q_i - p_iq_{i+1} = (-1)^i$. Moreover, let us define T_i so that $\alpha_{i-1}\alpha_i' = T_i + (-1)^{i+1}\omega_D$, i.e., we have $T_i = p_i(p_{i-1} - q_{i-1}) - q_iq_{i-1}\frac{D-1}{4}$ or $T_i = p_ip_{i-1} - Dq_iq_{i-1}$ when $D \equiv 1 \pmod{4}$ or $2, 3 \pmod 4$, respectively. Finally, set $c_i := [u_i, u_{i+1}, u_{i+2}, \dots]$ so that

$$c_i = u_i + \frac{1}{c_{i+1}}, \quad \omega_D = \frac{c_{i+1}p_i + p_{i-1}}{c_{i+1}q_i + q_{i-1}}. \tag{2.1}$$

Concerning norms of indecomposables, first note that as in the proof of [Ka2, Proposition 1] (or [JK, Theorem 4]), for odd i we have

$$N(\alpha_{i,r}) = \begin{cases} \frac{(D-1)/4 + (T_{i+1} - N_{i+1}r) - (T_{i+1} - N_{i+1}r)^2}{N_{i+1}}, & D \equiv 1 \pmod 4 \\ \frac{D - (T_{i+1} - N_{i+1}r)^2}{N_{i+1}}, & D \equiv 2, 3 \pmod 4. \end{cases} \tag{2.2}$$

In order to better estimate the sizes of norms, we will use the following two lemmas.

Lemma 4. *For $i \geq 0$ we have*

$$T_i = (-1)^i \left(\omega_D - \frac{N_{i-1}}{c_{i+1}} \right) \quad and \quad N_i = \frac{\sqrt{\Delta}}{c_{i+1}} - \frac{N_{i-1}}{c_{i+1}^2}. \tag{2.3}$$

Proof. As we indicated above, we will give the proof only in the case $D \equiv 1 \pmod 4$; for the other case see [Ka2, Proposition 5]. From the second formula in (2.1) we conclude $c_{i+1}\alpha_i' = -\alpha_{i-1}'$. Multiplying by $-\alpha_i$ we get

$$-N(\alpha_i)c_{i+1} = \alpha_i\alpha_{i-1}' = T_i + (-1)^{i+1}\omega_D'. \tag{2.4}$$

Moreover, we have

$$T_{i+1} = p_{i+1}(p_i - q_i) - q_iq_{i+1}\frac{D-1}{4}$$
$$= (u_{i+1}p_i + p_{i-1})(p_i - q_i) - (u_{i+1}q_i + q_{i-1})q_i\frac{D-1}{4}$$
$$= u_{i+1}N(\alpha_i) + T_i + p_iq_{i-1} - p_{i-1}q_i = u_{i+1}N(\alpha_i) + T_i + (-1)^{i+1}.$$

Plugging in (2.4), we obtain

$$(-1)^i\omega_D' - N(\alpha_{i+1})c_{i+2} = u_{i+1}N(\alpha_i) + (-1)^{i+1}\omega_D' - N(\alpha_i)c_{i+1} + (-1)^{i+1},$$

and so

$$N(\alpha_{i+1})c_{i+2} = (-1)^i(2\omega_D' + 1) + N(\alpha_i)(c_{i+1} - u_{i+1}) = (-1)^i\sqrt{D} + \frac{N(\alpha_i)}{c_{i+2}}.$$

This proves the formula for N_i in (2.3). The formula for T_i follows by plugging in the expression for N_i into (2.4). $\quad\square$

LEMMA 5. *For $i \geq 0$ we have*

$$\frac{\sqrt{\Delta}}{c_{i+1}}\left(1 - \frac{1}{c_i c_{i+1}}\right) < N_i < \frac{\sqrt{\Delta}}{c_{i+1}} \tag{2.5}$$

and in particular

$$\frac{1}{u_{i+1} + 10} < \frac{N_i}{\sqrt{\Delta}} \quad \textit{if } u_{i+1} \geq 3 \quad \textit{and} \quad \frac{N_i}{\sqrt{\Delta}} < \frac{1}{u_{i+1}}. \tag{2.6}$$

Proof. The upper bound in (2.5) follows immediately from the second formula in (2.3). The lower bound in the case $D \equiv 1 \pmod 4$ follows from

$$N_i = \frac{\sqrt{\Delta}}{c_{i+1}} - \frac{N_{i-1}}{c_{i+1}^2} = \frac{\sqrt{\Delta}}{c_{i+1}} - \frac{1}{c_{i+1}^2}\left(\frac{\sqrt{\Delta}}{c_i} - \frac{N_{i-2}}{c_i^2}\right) > \frac{\sqrt{\Delta}}{c_{i+1}}\left(1 - \frac{1}{c_i c_{i+1}}\right).$$

The bounds in (2.6) are simple consequences of (2.5) and the first formula in (2.1), cf. e.g. [Ka2, Theorem 8a]. $\qquad\square$

Finally, from [JW, Theorem 5.9] we conclude the following useful result. Note that by the discussion [JW, Section 5.2/5.3], principal reduced ideals correspond to convergents.

PROPOSITION 6. *Assume that $\mu \in \mathcal{O}_K$ is such that $0 < |N(\mu)| < \frac{1}{2}\sqrt{\Delta}$. Then $\mu = n\alpha_i$ or $\mu = n\alpha_i'$ for some $i \geq -1$ and $n \in \mathbb{Z}$.*

Note that Lemma 5 and Proposition 6 together give quite a close correspondence between primitive ideals of small norm and coefficients of the continued fraction: The elements α_i, $i = 0, 1, \ldots, s-1$ generate primitive (i.e., not divisible by a rational integer other than 1) principal ideals of norm $< \sqrt{\Delta}$. Conversely, the generators α of all primitive principal ideals α with $|N\alpha| < \sqrt{\Delta}/2$ come from a convergent to the continued fraction expansion of ω_D, so that the primitive principal ideals of norm $< \sqrt{\Delta}/2$ can naturally be injected into the set of $\{u_i \mid 1 \leq i \leq s\}$; the element α_i corresponds to the coefficient u_{i+1}.

Since the convergent α_j is totally positive if and only if j is odd, we also obtain an analogous correspondence between elements with negative norm modulo totally positive units and coefficients u_i with odd i, where $1 \leq i \leq s$ if s is even and $1 \leq i \leq 2s$ if s is odd. In terms of estimating the sum of coefficients u_i, we summarize this discussion as follows:

COROLLARY 7. a) *We have*

$$\sum_{N\mathfrak{a} < \sqrt{\Delta}/2}^{*} \frac{\sqrt{\Delta}}{N\mathfrak{a}} + O\Big(\sum_{N\mathfrak{a} < \sqrt{\Delta}}^{*} 1\Big) < \sum_{i=1}^{s} u_i < \sum_{N\mathfrak{a} < \sqrt{\Delta}}^{*} \frac{\sqrt{\Delta}}{N\mathfrak{a}},$$

where $\sum^{*}$ denotes the sum over principal primitive ideals.

b) *We have*

$$\sideset{}{^-}\sum_{N\mathfrak{a}<\sqrt{\Delta}/2} \frac{\sqrt{\Delta}}{N\mathfrak{a}} + O\Big(\sideset{}{^-}\sum_{N\mathfrak{a}<\sqrt{\Delta}} 1 \Big) < M_D < \sideset{}{^-}\sum_{N\mathfrak{a}<\sqrt{\Delta}} \frac{\sqrt{\Delta}}{N\mathfrak{a}},$$

where $\sum^-$ denotes the sum over principal primitive ideals generated by an element with negative norm.

Proof. The upper bound follows easily from the previous remark and the second bound in (2.6). For the lower bound we use the first bound in (2.6) to conclude that

$$\sum_{i=1}^{s} u_i > \sum_{\substack{i\le s \\ u_i\ge 3}} \left(\frac{\sqrt{\Delta}}{N_{i-1}} - 10 \right) \ge \sideset{}{^*}\sum_{N\mathfrak{a}<\sqrt{\Delta}/2} \frac{\sqrt{\Delta}}{N\mathfrak{a}} - 10 \sideset{}{^*}\sum_{N\mathfrak{a}<\sqrt{\Delta}} 1 - \sum_{\substack{i\le s \\ u_i\le 2}} \frac{\sqrt{\Delta}}{N_{i-1}}.$$

Using again the first bound in (2.6), we have $\sqrt{\Delta}/N_{i-1} \le u_i + 10$, so that we obtain

$$\sum_{i=1}^{s} u_i > \sideset{}{^*}\sum_{N\mathfrak{a}<\sqrt{\Delta}/2} \frac{\sqrt{\Delta}}{N\mathfrak{a}} - 22 \sideset{}{^*}\sum_{N\mathfrak{a}<\sqrt{\Delta}} 1$$

as desired. The lower bound in b) is proved similarly. $\square$

3   Upper bound

In this section we will prove an upper bound on the minimal number of variables of universal quadratic forms over $\mathbb{Q}(\sqrt{D})$ by constructing an explicit diagonal universal form that generalizes the results of Kim [Ki2]. The idea is to first express a totally positive integer as the sum of indecomposables. Since the set S_0 of indecomposables modulo (squares of) units is finite, we can collect together the terms corresponding to the same element of S_0, obtaining certain sums of units as coefficients.

Denote by $E = \mathbb{N}_0[\epsilon, \epsilon^{-1}]$ the semiring of all elements $\sum_{i=i_0}^{i_1} e_i \epsilon^i$ with $i_0, i_1 \in \mathbb{Z}$ and $e_i \in \mathbb{Z}$, $e_i \ge 0$ (recall that ϵ denotes the totally positive fundamental unit). Note that since ϵ is totally positive, we have that $E \subset \mathcal{O}_K^+ \cup \{0\}$.

PROPOSITION 8. *a) Every element $a \in \mathcal{O}_K^+$ is of the form $a = \sum_j n_j \sigma_j$ for some $n_j \ge 0$ and $\sigma_j \in S$.*
b) Every element $a \in \mathcal{O}_K^+$ is of the form $a = \sum_{\sigma \in S_0} e_\sigma \sigma$ for some $e_\sigma \in E$.

Proof. a) This is clear, as S is the set of all indecomposable elements (and there are no infinite chains $a_1 \succ a_2 \succ a_3 \succ \ldots$ of totally positive integers).
b) follows from a) by noting that if $\sigma \in S$, then there is some $k \in \mathbb{Z}$ such that $\sigma\epsilon^k \in S_0$. $\square$

We can now simplify the sums of units as follows. The proof of Lemma 9 is very similar to the proofs of Lemma 4 and Theorem 1 of [Ki2].

LEMMA 9. a) *For $e \in E$ there are $i, c, d \in \mathbb{Z}$ with $c, d \geq 0$ such that $e = c\epsilon^i + d\epsilon^{i+1}$.*
b) *All elements $e \in E$ are represented by the 8-ary form $x_1^2 + x_2^2 + x_3^2 + x_4^2 + \epsilon(x_5^2 + x_6^2 + x_7^2 + x_8^2)$.*

THEOREM 10. *The quadratic form*

$$\sum_{\sigma \in S_0} \sigma(x_{1,\sigma}^2 + x_{2,\sigma}^2 + x_{3,\sigma}^2 + x_{4,\sigma}^2 + \epsilon x_{5,\sigma}^2 + \epsilon x_{6,\sigma}^2 + \epsilon x_{7,\sigma}^2 + \epsilon x_{8,\sigma}^2)$$

is universal and has $8M_D$ variables.

Proof. The universality of the form follows by Proposition 8 and Lemma 9. Its number of variables is eight times the number of elements of S_0.
When s is even, the elements of S_0 are of the form $\alpha_i + r\alpha_{i+1}$ with odd i, $-1 \leq i \leq s-3$ and $0 \leq r < u_{i+2}$. The same is true when s is odd, except that we take $-1 \leq i \leq 2s-3$, so that $|S_0| = M_D$. $\qquad\square$

Finally, let us note that this result is exactly analogous to Kim's case of $D = n^2 - 1$ [Ki2]. Then we have $\sqrt{D} = [n-1, \overline{1, 2(n-1)}]$, and so $s = 2$, $u_1 = 1$, and $S_0 = \{1\}$, also obtaining $M_D = 1$ and $m_{\text{diag}}(D) \leq 8$.

4   LOWER BOUND

Let us now prove a lower bound on the number of variables of a diagonal universal quadratic form. Assume that $Q(x_i) = \sum_{1 \leq i \leq m} a_i x_i^2$ is a universal totally positive diagonal quadratic form with $a_i \in \mathcal{O}_K^+$. For every indecomposable element σ we know that $\sigma = Q(x_i)$ for some $x_1, \ldots, x_m$, but since σ is indecomposable, this is possible only when $\sigma = a_i x_i^2$ for some i. Hence it will be important for us to understand when it can happen that squares of elements of $\mathcal{O}_K$ are indecomposable.

LEMMA 11. *Assume that $\alpha \in \mathcal{O}_K$ is such that α^2 is indecomposable. Then for some $j \geq -1$ we have $\alpha \in \{\pm\alpha_j, \pm\alpha_j'\}$.*

Proof. By the theorem of Dress-Scharlau [DS, Theorem 3] we know that $N(\alpha^2) \leq \Delta/4$, and so $|N(\alpha)| \leq \frac{1}{2}\sqrt{\Delta}$. Since α must obviously be primitive, the claim follows from Proposition 6. $\qquad\square$

Using the previous lemma, we are now ready to prove a lower bound on the number of variables.

THEOREM 12. *Every diagonal universal totally positive quadratic form over $\mathcal{O}_K$ needs at least M_D/s variables if s is even and $M_D/2s$ if s is odd.*

Proof. Let $T = S_0$ if s is odd and $T = \{\sigma \in S \mid \epsilon^2 > \sigma \geq \sigma' > 0\}$ if s is even, i.e., we can describe T uniformly as $T = \{\sigma \in S \mid \alpha_{s-1}^2 = \alpha_{2s-1} > \sigma \geq \sigma' > 0\}$. Thus T has cardinality $2M_D$ if s is even and M_D if s is odd.

Note that if $1 \neq \sigma = x - y\omega'_D \in T$, then $x, y > 0$ and $\sigma > 1 > \sigma'$.

As before, let $Q(x_i) = \sum_{1 \leq i \leq m} a_i x_i^2$ be a universal form with $a_i \in \mathcal{O}_K^+$. Multiplying any of the variables by a power of the unit α_{s-1} is an invertible substitution, and so we can assume without loss of generality that the coefficients a_i satisfy $0 < a_i < a'_i < \alpha_{s-1}^2$.

The only way in which Q can represent an indecomposable σ is when all but one of the variables x_i are zero. Hence for each $\sigma \in T$ we have $\sigma = a_{i(\sigma)} \alpha(\sigma)^2$ for some index $1 \leq i(\sigma) \leq m$ and some $\alpha(\sigma) \in \mathcal{O}_K$, which must be indecomposable. By Lemma 11, $\alpha(\sigma)$ is, up to sign, a convergent α_j or its conjugate α'_j for some j. Moreover, we have the two bounds

$$|\alpha(\sigma)| = (\sigma/a_{i(\sigma)})^{1/2} \begin{cases} < \alpha_{s-1} \cdot (a'_{i(\sigma)})^{1/2} < \alpha_{s-1}^2, \\ > (\sigma'/a'_{i(\sigma)})^{1/2} = |\alpha(\sigma)'|. \end{cases}$$

In particular, this excludes the possibility $\alpha(\sigma) = \pm\alpha'_j$ and forces $\alpha(\sigma) = \pm\alpha_j$ with $-1 \leq j \leq 2s - 2$.

Since all elements of T are represented by Q, we have $T \subset \{a_i \alpha_j^2 \mid 1 \leq i \leq m, -1 \leq j \leq 2s - 2\}$. Comparing cardinalities, we get $\#T \leq m \cdot 2s$, finishing the proof. $\square$

Comparing this bound on the number of variables with Theorem 10, we see that it is asymptotically good when s is small. This of course happens only rarely, and so it would be very interesting to be able to replace $1/s$ in Theorem 12 by an absolute constant.

This is perhaps hard, but we can almost achieve this, at the cost of having to consider the sum $M_{D,\varepsilon}^*$ of sufficiently large coefficients u_j instead of M_D, as we are now going to prove.

Let us first consider in more detail the set of semi-convergents $\alpha_{i,r}$ with fixed odd i and varying r. Their number is $u_{i+2} \asymp \sqrt{D}/N_{i+1}$, and so if N_{i+1} is small, there are many of them. Let $N := N_{i+1}$, $T := T_{i+1}$ and $u := u_{i+2}$ so that

$$N(\alpha_{i,r}) = \begin{cases} \frac{D - (T - Nr)^2}{N}, & D \equiv 2, 3 \pmod{4}, \\ \frac{(D-1)/4 + (T - Nr) - (T - Nr)^2}{N}, & D \equiv 1 \pmod{4} \end{cases}$$

by (2.2). In either case, we view this as a quadratic polynomial $f(r)$ and note that from the definition of $\alpha_{i,r}$ as (totally positive) indecomposables we have

$$\mathbb{Z} \cap [0, u] = \{n \in \mathbb{Z} \mid f(n) > 0\}$$

(using the estimate for T from (2.3), this also corresponds to $\left|\frac{T}{N} - \frac{u}{2}\right|$ being small). We define the multiplicative arithmetic function $\rho_f(d) = \{n \pmod{d} \mid f(n) \equiv 0 \pmod{d}\}$. An application of Hensel's lemma shows

$$\rho_f(p^k) \leq 2 \tag{4.1}$$

for all primes p and all $k \geq 2$:

We verify this first in the case $f(x) = \frac{D-(T-Nx)^2}{N} = \frac{D-T^2}{N} + 2Tx - Nx^2$.

If $p \mid D$, then $D \equiv (T - Nx)^2 \pmod{p^2}$ has no solution since D is squarefree, i.e., $\rho(p^k) = 0$ for all $k \geq 2$.

If $p \nmid D$ and $p \nmid N$, then N is invertible modulo p, and so $\rho(p) \leq 2$ and if $p \neq 2$, we can use Hensel's lemma to conclude $\rho(p^k) \leq 2$ for every $k \geq 1$. If $p = 2$, then we see directly that $\rho(2^k) = 0$ for all $k \geq 2$, since $D \equiv 2, 3 \not\equiv 1 \equiv (T - Nx)^2$ (mod 4).

If $p \nmid 4D$ and $p \mid N$, then $f(x) \pmod p$ is linear or constant, and so $\rho(p) \leq 1$ and $\rho(p^k) \leq 1$ for every $k \geq 1$ by Hensel's lemma.

Finally, if $p = 2 \nmid D$ and $2 \mid N$, then $f(x) \equiv \frac{D-T^2}{N} \pmod 2$. Since $D \equiv 2, 3$ (mod 4), we have $4 \nmid D - T^2$ and we see that $\rho(2^k) = 0$ for all $k \geq 1$ in this case.

In the second case we have

$$f(x) = \frac{D - (2(T - Nx) - 1)^2}{4N} = \frac{D - (2T - 1)^2}{4N} + (2T - 1)x - Nx^2.$$

The discussion is similar to the previous one:

If $p \mid D$, then again $D \equiv (2(T - Nx) - 1)^2 \pmod{p^2}$ has no solution, i.e., $\rho(p^k) = 0$ for all $k \geq 2$.

If $p \nmid D$ and $p \nmid N$, then N is invertible modulo p, and so $\rho(p) \leq 2$ and if $p \neq 2$, we can use Hensel's lemma to conclude $\rho(p^k) \leq 2$ for every $k \geq 1$. For $p = 2$ we have that $f(x)$ is constant modulo 2 and non-constant modulo 4, and so $\rho(2^k) \leq 2$ for $k \geq 2$.

If $p \nmid 4D$ and $p \mid N$, then $f(x) \pmod p$ is linear or constant, and so $\rho(p) \leq 1$ and $\rho(p^k) \leq 1$ for every $k \geq 1$ by Hensel's lemma.

If $p = 2 \nmid D$ and $2 \mid N$, then $f(x) \equiv \frac{D-(2T-1)^2}{4N} - x \pmod 2$, and so $\rho(2) = 1$. Distinguishing some more cases, one gets $\rho(4) = 1$ and $\rho(2^k) \leq 2$ for $k \geq 3$.

Having established (4.1), we can now prove the following lemma.

LEMMA 13. *Let f, T, N, u be as above. Let $k \in 2\mathbb{N}$ and $1 \leq X \leq u$. Then*

$$\mathcal{S} := \#\{n \leq X \mid f(n) \text{ is } k\text{-th power free}\}$$

$$= C_{k,f} X + O\left(Y X^\varepsilon + \frac{X}{Y^{k-1}} + \frac{D^{1+\varepsilon}}{N Y^k}\right)$$

for any $\varepsilon > 0$, where $C_{k,f} \geq 1/\zeta(k)^3$, and $1 \leq Y \leq X$ can be chosen arbitrarily.

Proof. For $1 \leq Y \leq X$ we have

$$\mathcal{S} = \sum_{n \leq X} \sum_{d^k \mid f(n)} \mu(d) = \sum_{d \leq Y} \mu(d) \sum_{\substack{n \leq X \\ d^k \mid f(n)}} 1 + \sum_{\substack{n \leq X \\ f(n) = d^k m \\ d > Y}} \mu(d). \qquad (4.2)$$

The first term on the right hand side equals

$$\sum_{d \leq Y} \mu(d) \left(\frac{\rho_f(d^k)}{d^k} X + O(X^\varepsilon)\right) = C_{k,f} X + O(Y X^\varepsilon + Y^{-k+1} X)$$

with

$$C_{k,f} = \sum_{d=1}^{\infty} \mu(d)\frac{\rho_f(d^k)}{d^k} = \prod_p \left(1 - \frac{\rho_f(p^k)}{p^k}\right) \geq \prod_p \left(1 - \frac{2}{p^k}\right) \geq \frac{1}{\zeta(k)^3}.$$

For the second term, we observe that $0 < f(n) \leq D/N$ for $1 \leq n \leq X$, so that $0 < m \leq D/(NY^k)$. The summation condition $f(n) = d^k m$ is equivalent to the equation $(d^{k/2})^2 mN + (T - Nn)^2 = D$ resp. $4(d^{k/2})^2 mN + (2(T - Nn) - 1)^2 = D$. For given m, N, T, D, this has at most $6\tau(D)$ solutions (d, n), since the imaginary quadratic number field $\mathbb{Q}(\sqrt{-mN})$ can have at most 6 units and the number of principal ideals of norm D is at most $\tau(D) \ll D^\varepsilon$. Hence we can bound the second term in (4.2) by $O(D^{1+\varepsilon}Y^{-k}N^{-1})$. □

Now we are ready to bound the minimal number of variables from below in terms of the sum $M_{D,\varepsilon}^*$ of coefficients $u_i \geq D^{1/8+\varepsilon}$, defined immediately before Theorem 1.

THEOREM 14. *We have*

$$m_{\mathrm{diag}}(D) \gg M_{D,\varepsilon}^*,$$

where the implicit constant depends only on $\varepsilon > 0$.

Proof. Let's consider the (pairwise distinct) indecomposables $\alpha_{i,r}$ with odd i and $0 \leq r < u$. Arguing similarly as in the proof of Theorem 12, each of them has to be of the form $\sigma = a_{j(\sigma)}\alpha(\sigma)^2$, where a_j are the coefficients of a universal form. Hence it suffices to show that for a positive density of suitable values of i, r we have that the elements $\alpha_{i,r}$ have 4th-power-free norms (because then the indices $j(\alpha_{i,r})$ must be different for different values of i, r). Note that for $-1 \leq i \leq 2s-3$, the indecomposables $\alpha_{i,r}$ are pairwise distinct modulo squares of units, and so we can consider the cases of different i's separately and then just add them together.

Hence let i be odd. As before we write $u := u_{i+2}$, $N := N_{i+1}$, and $T := T_{i+1}$. We apply the previous lemma with $X = u$, $Y = u^{1-2\varepsilon}$ and $k = 4$. Provided $N \leq D^{3/8-\varepsilon}$ (for $\varepsilon > 0$ sufficiently small and D sufficiently large) we can conclude from (2.6) that $u \asymp D^{1/2}N^{-1}$ and

$$\mathcal{S} \gg u + O(u^{1-\varepsilon} + D^{1+\varepsilon}N^{-1}u^{-4+8\varepsilon}) = u + O\left(u^{1-\varepsilon}(1 + N^{4-9\varepsilon}D^{-\frac{3}{2}+\frac{11}{2}\varepsilon})\right) \gg u$$

as desired. □

5   ANALYTIC PREPARATION

We now employ analytic methods to prove Theorems 2 and 3. We denote by $\mathcal{C}$ the class group of K (of order h) and by $\widehat{\mathcal{C}}$ the corresponding character group. For each $\chi \in \widehat{\mathcal{C}}$ we denote by $L_K(s, \chi)$ the corresponding class group L-function. It is entire except if $\chi = \chi_0$ is the trivial character, in which case

$L_K(s, \chi_0) = \zeta(s)L(s, \chi_\Delta)$ has a simple pole at $s = 1$ with residue $L(1, \chi_\Delta)$. We have the important subconvexity bound

$$L_K(1/2 + it, \chi) \ll (1 + |t|)^A D^{\frac{1}{4} - \delta} \tag{5.1}$$

where $A, \delta > 0$ are some absolute constants. The first such result was obtained in [DFI], the best value of $\delta = 1/1889$ is due to [BHM]. We write

$$
\begin{aligned}
L_K^*(s, \chi) &:= \sum_{\mathfrak{a} \text{ primitive}} \frac{\chi(\mathfrak{a})}{(N\mathfrak{a})^s} \\
&= \prod_{p = \mathfrak{p}\bar{\mathfrak{p}} \text{ split}} \left(1 + \sum_{n=1}^{\infty} \frac{2\Re\chi(\mathfrak{p})}{p^{ns}}\right) \prod_{p = \mathfrak{p}^2 \text{ ramified}} \left(1 - \frac{\chi(\mathfrak{p})}{p^s}\right)^{-1} \\
&= L_K(s, \chi)\zeta^{(\Delta)}(2s)^{-1}
\end{aligned}
$$

where the superscript (Δ) denotes the removal of Euler factors at primes dividing Δ.

For $s > 1$ we have

$$\frac{1}{h} \sum_{\chi \neq \chi_0} L_K(s, \chi) = \zeta(s, \text{princ}) - \frac{1}{h}\zeta(s)L(s, \chi_\Delta),$$

where

$$\zeta(s, \text{princ}) = \sum_{\mathfrak{a} \text{ principal}} \frac{1}{(N\mathfrak{a})^s}$$

is the zeta-function associated to the class of principal ideal. If we write

$$L(D) = \lim_{s \to 1} \left[\zeta(s, \text{princ}) - \frac{L(1, \chi_\Delta)/h}{s - 1}\right], \tag{5.2}$$

then comparing Taylor coefficients in the equation

$$\frac{1}{h} \sum_{\chi \neq \chi_0} L_K(s, \chi) = \zeta(s, \text{princ}) - \frac{1}{h}L(s, \chi_\Delta)\zeta(s)$$

we obtain

$$\frac{1}{h} \sum_{\chi \neq \chi_0} L_K(1, \chi) = L(D) - \frac{\gamma L(1, \chi_\Delta) + L'(1, \chi_\Delta)}{h},$$

where $\gamma = 0.577\ldots$ denotes Euler's constant, and so

$$\frac{1}{h} \sum_{\chi \neq \chi_0} L_K^*(1, \chi) = \frac{1}{\zeta^{(\Delta)}(2)} \left(L(D) - \frac{\gamma L(1, \chi_\Delta) + L'(1, \chi_\Delta)}{h}\right). \tag{5.3}$$

Explicit, but very complicated expressions for $L(D)$ have been obtained by Hecke [Hec], Herglotz [Her], and Zagier [Za1] and can be regarded as a real-quadratic analogue of Kronecker's limit formula.

In the following we fix a smooth function $w : [0, \infty) \to [0, 1]$ that is 1 on $[0, 1]$ and 0 on $[2, \infty)$. Its Mellin transform $\widehat{w}(s) = \int_0^\infty w(x) x^{s-1} dx$ (initially in $\Re s > 0$) is entire except for a simple pole at $s = 0$ with residue 1, which can be seen by decomposing $w = w|_{[0,1]} + w|_{(1,2]}$. More precisely, it has a Laurent expansion

$$\widehat{w}(s) = \frac{1}{s} + c + \dots, \quad c = \int_1^2 w(x) \frac{dx}{x}.$$

Moreover, it is rapidly decaying on vertical lines:

$$\widehat{w}(s) \ll_B |s|^{-B} \tag{5.4}$$

for any constant $B \geq 1$ and $\Re s > -1$ (say), as can be seen by partial integration.

LEMMA 15. *Let $X \geq 1$. Then*

$$\sum_{\mathfrak{a}}^* \frac{1}{N\mathfrak{a}} w\left(\frac{N\mathfrak{a}}{X}\right) = \frac{1}{\zeta^{(\Delta)}(2)} \left(L(D) + \frac{L(1, \chi_\Delta)(\log X + c + (\log \zeta^{(\Delta)})'(2))}{h} \right)$$
$$+ O\left(X^{-\frac{1}{2}} D^{\frac{1}{4} - \frac{1}{2000}} \right),$$

and

$$\sum_{\mathfrak{a}}^* w\left(\frac{N\mathfrak{a}}{X}\right) = \frac{\widehat{w}(1)}{\zeta^{(\Delta)}(2)} \frac{X}{h} L(1, \chi_\Delta) + O(X^{\frac{1}{2}} D^{\frac{1}{4} - \frac{1}{2000}}),$$

where as before $\sum^*$ denotes a sum over primitive principal ideals and $(\log \zeta^{(\Delta)})'(2) := \frac{d}{ds}(\log \zeta^{(\Delta)}(s))\big|_{s=2}$.

Proof. This is a standard contour shift argument. We have by Mellin inversion and orthogonality of class group characters

$$\sum_{\mathfrak{a}}^* \frac{1}{N\mathfrak{a}} w\left(\frac{N\mathfrak{a}}{X}\right) = \int_{(2)} \frac{1}{h} \sum_{\chi \in \widehat{\mathcal{C}}} L_K^*(s+1, \chi) \widehat{w}(s) X^s \frac{ds}{2\pi i}$$

where here and in the following $\int_{(c)}$ denotes a complex contour integral over the vertical line $\Re s = c$. Shifting the contour to $\Re s = -1/2$, computing the residue of the double pole at $s = 0$, namely

$$\frac{1}{h} \sum_{\chi \neq \chi_0} L_K^*(1, \chi) + \frac{L(1, \chi_\Delta)(\log X + c + (\log \zeta^{(\Delta)})'(2) + \gamma) + L'(1, \chi_\Delta)}{\zeta^{(\Delta)}(2)h}$$
$$= \frac{1}{\zeta^{(\Delta)}(2)} \left(L(D) + \frac{L(1, \chi_\Delta)(\log X + c + (\log \zeta^{(\Delta)})'(2))}{h} \right) \tag{5.5}$$

by (5.3), and estimating the remaining integral with (5.1) and (5.4) gives the first formula. The proof of the second formula is similar. $\square$

The following lemma is an application of Burgess' bound.

LEMMA 16. *For $\varepsilon > 0$ we have*

$$L'(1, \chi_\Delta) \geq -\left(\frac{3}{8} + \varepsilon\right) L(1, \chi_\Delta) \left(\log D + O_\varepsilon(1)\right).$$

Proof. Let

$$r(n) := \sum_{d|n} \chi_\Delta(d).$$

For $X \geq 1$ we compute by Mellin inversion

$$\mathcal{S} := \sum_{n \leq X} \frac{r(n)}{n} \left(1 - \frac{n}{X}\right) = \int_{(2)} \zeta(s+1) L(s+1, \chi_\Delta) \frac{X^s}{s(s+1)} \frac{ds}{2\pi i}.$$

We shift the contour to $\Re s = -1/2$. The residue at $s = 0$ contributes

$$L(1, \chi_\Delta)(\log X + \gamma - 1) + L'(1, \chi_\Delta), \tag{5.6}$$

while Heath-Brown's hybrid bound [HB]

$$L(1/2 + it, \chi_\Delta) \ll ((1 + |t|)D)^{3/16 + \varepsilon} \tag{5.7}$$

along with the classical convexity estimate $\zeta(1/2 + it) \ll (1 + |t|)^{1/4 + \varepsilon}$ bounds the remaining integral by

$$O(X^{-1/2} D^{3/16 + \varepsilon}). \tag{5.8}$$

On the other hand, for a parameter $1 \leq Y \leq X/10$ we have by partial summation

$$\mathcal{S} \geq \sum_{Y \leq n \leq X} \frac{r(n)}{n} \left(1 - \frac{n}{X}\right) = \int_Y^X \sum_{Y \leq n \leq t} r(n) \frac{dt}{t^2} \geq \int_{3Y}^X \sum_{Y \leq n \leq t} r(n) \frac{dt}{t^2}.$$

For $0 < \varepsilon < 1/10$ and $t \geq 3Y$ let $V_t : \mathbb{R} \to [0, 1]$ be a smooth function with support in $[Y, t]$ that is 1 on $[2Y, (1 - \varepsilon)t]$ and satisfies $V_t^{(j)}(x) \ll_{\varepsilon, j} x^{-j}$ for $j \in \mathbb{N}_0$ uniformly in t and Y. It follows from the definition and repeated integration by parts that

$$\widehat{V_t}(s) = \int_0^\infty V_t(s) x^{s-1} dx \ll t^{\Re s} |s|^{-10} \tag{5.9}$$

for $\Re s \geq 1/2$. Then again by Mellin inversion we have

$$\mathcal{S} \geq \int_{3Y}^X \sum_n r(n) V_t(n) \frac{dt}{t^2} = \int_{3Y}^X \int_{(2)} \zeta(s) L(s, \chi_\Delta) \widehat{V_t}(s) \frac{ds}{2\pi i} \frac{dt}{t^2}.$$

Shifting the s-contour to $\Re s = 1/2$ and using (5.7) and (5.9), we obtain

$$\mathcal{S} \geq \int_{3Y}^{X} \left(\left(t(1-\varepsilon) - 2Y \right) L(1, \chi_\Delta) + O(t^{1/2} D^{3/16+\varepsilon}) \right) \frac{dt}{t^2}$$

$$= L(1, \chi_\Delta) \left((1-\varepsilon) \log \frac{X}{3Y} + O(1) \right) + O(D^{3/16+\varepsilon} Y^{-1/2}).$$

Combining this with (5.6) and (5.8) yields

$$L(1, \chi_\Delta)(\log X + O(1)) + L'(1, \chi_\Delta)$$

$$\geq L(1, \chi_\Delta) \left((1-\varepsilon) \log \frac{X}{Y} + O(1) \right) + O(D^{3/16+\varepsilon} Y^{-1/2}).$$

Now we choose

$$Y = D^{3/8+4\varepsilon}, \quad X = D \text{ (say)}$$

getting

$$L'(1, \chi_\Delta) \geq - \left(\frac{3}{8} + 5\varepsilon \right) L(1, \chi_\Delta) \left(\log D + O(1) \right) + O(D^{-\varepsilon}).$$

By Siegel's theorem, the last error term is negligible compared to the main term, and the result follows. $\square$

6 BOUNDS FOR THE SUM OF CONVERGENTS

THEOREM 17. *We have*

$$\sum_{i=1}^{s} u_i = \frac{\sqrt{\Delta}}{\zeta^{(\Delta)}(2)} \Big(L(D) + \frac{1}{h} L(1, \chi_\Delta)(\log \sqrt{D} + O(1)) \Big) + O(D^{\frac{1}{2} - \frac{1}{2000}}).$$

Proof. Using Corollary 7 we prove an upper bound and a lower bound. On the one hand we have

$$\sum_{i=1}^{s} u_i \leq \sideset{}{^*}\sum_{N\mathfrak{a} \leq \sqrt{\Delta}} \frac{\sqrt{\Delta}}{N\mathfrak{a}} \leq \sideset{}{^*}\sum_{\mathfrak{a}} \frac{\sqrt{\Delta}}{N\mathfrak{a}} w \left(\frac{N\mathfrak{a}}{\sqrt{\Delta}} \right),$$

so that

$$\sum_{i=1}^{s} u_i \leq \frac{\sqrt{\Delta}}{\zeta^{(\Delta)}(2)} \Big(L(D) + \frac{1}{h} L(1, \chi_\Delta)(\log \sqrt{D} + O(1)) \Big) + O(D^{\frac{1}{2} - \frac{1}{2000}})$$

by Lemma 15. On the other hand, since $\text{supp}(w) \subseteq [0, 2]$, we have

$$\sum_{i=1}^{s} u_i \geq \sideset{}{^*}\sum_{\mathfrak{a}} \frac{\sqrt{\Delta}}{N\mathfrak{a}} w \left(\frac{2N\mathfrak{a}}{\frac{1}{2}\sqrt{\Delta}} \right) + O \left(\sideset{}{^*}\sum_{\mathfrak{a}} w \left(\frac{N\mathfrak{a}}{\sqrt{\Delta}} \right) \right),$$

getting a corresponding lower bound again by Lemma 15. $\square$

Since $u_s = 2\lfloor \omega_D \rfloor$ or $2\lfloor \omega_D \rfloor - 1$, it is clear that the error term $O(D^{\frac{1}{2} - \frac{1}{2000}})$ is smaller than the main term. However, since $L(D)$ may be negative, it is not clear if $L(D) + \log \sqrt{D} L(1, \chi_\Delta)/h$ grows more quickly than $L(1, \chi_{-D})/h$ as $D \to \infty$, so that a priori Theorem 17 may only give an upper bound. In order to understand the behaviour of the main term, it is more convenient to use the left hand side of the expression (5.5) with $X = \sqrt{D}$. Notice that all L-values are real (since $L(s, \chi) = L(s, \bar{\chi}) = \overline{L(s, \chi)}$ for $\chi \in \widehat{\mathcal{C}}$ and $s \in \mathbb{R}$) and, except for $L'(1, \chi_\Delta)$, positive and that $(\log \zeta^{(\Delta)})'(2) \asymp 1$. Since Lemma 16 implies

$$L(1, \chi_\Delta) \log \sqrt{D} + L'(1, \chi_\Delta) \gg L(1, \chi_\Delta) \log \sqrt{D},$$

for sufficiently large D, (5.5) implies

$$L(D) + \frac{1}{h} L(1, \chi_\Delta)(\log \sqrt{D} + O(1)) \gg \frac{1}{h} L(1, \chi_\Delta) \log \sqrt{D}, \qquad (6.1)$$

for sufficiently large D, and hence

Corollary 18. *We have*

$$\sum_{i=1}^{s} u_i \sim \frac{\sqrt{\Delta}}{\zeta^{(\Delta)}(2)} \Big(L(D) + \frac{1}{h} L(1, \chi_\Delta) \log \sqrt{D} \Big)$$

as $D \to \infty$.

We have the classical upper bounds [Fo, Lemma 4]

$$L(s, \chi_\Delta) \ll \log \Delta, \quad L_K(s, \chi) \ll (\log \Delta)^2 \qquad (6.2)$$

for $|s - 1| \le 1/\log \Delta$ and $\chi \in \widehat{\mathcal{C}}$, from which we also conclude $L'(1, \chi_\Delta) \ll (\log \Delta)^2$ by Cauchy's integral formula. Thus we obtain

$$\sqrt{\Delta} \le 2\lfloor \omega_D \rfloor \le \sum_{i=1}^{s} u_i \ll \sqrt{\Delta}(\log \Delta)^2$$

from (5.5), which by (1.2) and (1.3) implies the upper bound in Theorem 2. As mentioned in the introduction, the lower bound $M_{D,\varepsilon}^* \ge \sqrt{\Delta}$ is trivial if s is odd.

Concluding remarks:
(1) If one is willing to assume the Generalized Riemann Hypothesis, then $L(1, \chi_\Delta) \gg 1/\log \log D$ [Lit], and instead of (6.2) we have

$$L(1, \chi_\Delta), L_K(1, \chi) \ll \log \log D, \quad L'(1, \chi_\Delta) \ll (\log \log D)^2$$

(cf. [Li, Section 5]), so that again by (5.5) for the upper bound and (6.1) for the lower bound we obtain

$$\sqrt{D} \left(1 + \frac{\log D}{h \log \log D} \right) \ll \sum_{i=1}^{s} u_i \ll \sqrt{D} \left(1 + \frac{\log D}{h} \right) (\log \log D)^2. \qquad (6.3)$$

(2) Using the second part of Lemma 15, we obtain easily

$$\frac{\sqrt{D}}{h}L(1,\chi_\Delta) + O(D^{\frac{1}{2}-\frac{1}{2000}}) \ll s \ll \frac{\sqrt{D}}{h}L(1,\chi_\Delta) + O(D^{\frac{1}{2}-\frac{1}{2000}}), \qquad (6.4)$$

where the lower bound is of course a trivial statement if $h \geq D^{1/2000}$ (but conjecturally this does not happen very frequently).

(3) In the same way as in Theorem 17 one can show

$$M_D = \frac{\sqrt{\Delta}}{\zeta^{(\Delta)}(2)}\Big(L^-(D) + \frac{1}{h^+}L(1,\chi_\Delta)(\log\sqrt{D} + O(1))\Big) + O(D^{\frac{1}{2}-\frac{1}{2000}})$$

where h^+ is the narrow class number (which equals h if $\mathcal{O}_K$ has a unit of negative norm and otherwise equals $2h$) and $L^-(D)$ is the constant Taylor coefficient of $\zeta(s,\mathrm{princ}^-) = \sum_{\mathfrak{a}}(N\mathfrak{a})^{-s}$. Alternatively, one can also combine Corollary 18 with [Za2, Satz 2, §14] for the alternating sum, to obtain a formula form for M_D. In many cases, however, the main term will be smaller than the error term, so that it is unclear to which extent this formula is meaningful.

References

[Bh] M. Bhargava, *On the Conway-Schneeberger Fifteen Theorem*, Contemp. Math. 272 (1999), 27-37

[BH] M. Bhargava, J. Hanke, *Universal quadratic forms and the 290-theorem*, preprint

[BG] A. Biró, A. Granville, *Zeta functions for ideal classes in real quadratic fields, at $s = 0$*, J. Number Theory 132 (2012), 1807-1829

[BHM] V. Blomer, G. Harcos, P. Michel, *Bounds for modular L-functions in the level aspect*, Ann. Sci. Ecole Norm. Sup. 40 (2007), 697-740

[BK] V. Blomer, V. Kala, *Number fields without universal n-ary quadratic forms*, Math. Proc. Cambridge Philos. Soc. 159 (2015), 239-252

[CKR] W. K. Chan, M.-H. Kim, S. Raghavan, *Ternary universal integral quadratic forms*, Japan. J. Math. 22 (1996), 263-273

[DS] A. Dress, R. Scharlau, *Indecomposable totally positive numbers in real quadratic orders*, J. Number Theory 14 (1982), 292-306

[DFI] W. Duke, J. Friedlander, H. Iwaniec, *The subconvexity problem for Artin L-functions*, Invent. Math. 149 (2002), 489-577

[Fr] C. Friesen, *On continued fractions of given period*, Proc. Amer. Math. Soc. 103 (1988), 8-14

[Fo] E. Fogels, *On the zeros of Hecke's L-functions I*, Acta Arith. 7 (1962), 87-106

[HW] G. H. Hardy, E. M. Wright, *An introduction to the theory of numbers*, 5th edition. The Clarendon Press, Oxford University Press, New York, 1979

[HB] Heath-Brown, *Hybrid bounds for Dirichlet L-functions. II*, Quart. J. Math. Oxford 31 (1980), 157-167

[Hec] E. Hecke, *Über die Kroneckersche Grenzformel für reelle quadratische Körper und die Klassenzahl relativ-abelscher Körper*, Verhandl. d. Naturforschenden Gesell. Basel 28, 363-372 (1917). (Mathematische Werke pp.198-207)

[Her] G. Herglotz, *Über die Kroneckersche Grenzformel für reelle, quadratische Körper I, II*, Berichte Verhandl. Sächs. Akad. Wiss. Leipzig 75 (1923), 3-14, 31-37

[JK] S. W. Jang, B. M. Kim, *A refinement of the Dress-Scharlau theorem*, J. Number Theory 158 (2016), 234-243

[JW] M. Jacobson, H. Williams, *Solving the Pell equation*, CMS books in mathematics, Springer-Verlag 2009

[Ka1] V. Kala, *Universal quadratic forms and elements of small norm in real quadratic fields*, Bull. Aust. Math. Soc. 94 (2016), 7-14

[Ka2] V. Kala, *Norms of indecomposable integers in real quadratic fields*, J. Number Theory 166 (2016), 193-207

[Ki1] B. M. Kim, *Finiteness of real quadratic fields which admit positive integral diagonal septenary universal forms*, Manuscr. Math. 99 (1999), 181-184

[Ki2] B. M. Kim, *Universal octonary diagonal forms over some real quadratic fields*, Commentarii Math. Helv. 75 (2000), 410-414

[Li] X. Li, *Upper bounds on L-functions at the edge of the critical strip*, IMRN 2010, 727-755

[Lit] J. E. Littlewood, *On the class-number of the corpus $P(\sqrt{-k})$*, Proc. Lond. Math. Soc 27 (1928), 358-372

[Ma] H. Maaß, *Über die Darstellung total positiver Zahlen des Körpers $R(\sqrt{5})$ als Summe von drei Quadraten*, Abh. Math. Sem. Univ. Hamburg 14 (1942), 185-191

[Pe] O. Perron, *Die Lehre von den Kettenbrüchen*, Chelsea, New York, 1958

[Sa]   H. Sasaki, *Quaternary universal forms over $\mathbb{Q}[\sqrt{13}]$*, Ramanujan J. 18 (2009), 73-80

[Si]   C. L. Siegel, *Sums of m-th powers of algebraic integers*, Ann. Math. 46 (1945), 313-339

[Ya]   P. Yatsyna, *A lower bound for the rank of a universal quadratic form with integer coefficients in a totally real field*, Comment. Math. Helv., to appear

[Za1]   D. Zagier, *A Kronecker limit formula for real quadratic fields*, Math. Ann. 213 (1975), 153-184

[Za2]   D. Zagier, *Zetafunktionen und quadratische Körper*, Springer-Verlag 1981

Valentin Blomer
University of Göttingen
Mathematisches Institut
Bunsenstr. 3-5
D-37073 Göttingen
Germany
vblomer@math.uni-goettingen.de

Vítězslav Kala
University of Göttingen
Mathematisches Institut
Bunsenstr. 3-5
D-37073 Göttingen
Germany
and
Charles University
Faculty of Mathematics and
Physics
Department of Algebra
Sokolovská 83
18600 Praha 8
Czech Republic
vita.kala@gmail.com