ON WARING'S PROBLEM IN NUMBER FIELDS

MORLEY DAVIDSON

1. *Introduction*

Let $\mathbb{K}$ be an algebraic number field of degree n over the rationals, and denote by J_k the subring of $\mathbb{K}$ generated by the kth powers of the integers of $\mathbb{K}$. Then $G_{\mathbb{K}}(k)$ is defined to be the smallest $s \geqslant 1$ such that, for all totally positive integers $\nu \in J_k$ of sufficiently large norm, the Diophantine equation

$$\nu = \lambda_1^k + \ldots + \lambda_s^k \tag{1.1}$$

is soluble in totally non-negative integers λ_i of $\mathbb{K}$ satisfying

$$N(\lambda_i) \ll N(\nu)^{1/k} \quad (1 \leqslant i \leqslant s). \tag{1.2}$$

In (1.2) and throughout this paper, all implicit constants are assumed to depend only on $\mathbb{K}$, k, and s. The notation $G_{\mathbb{K}}(k)$ generalizes the familiar symbol $G(k)$ used in Waring's problem, since we have $G_{\mathbb{Q}}(k) = G(k)$.

By extending the Hardy–Littlewood circle method to number fields, Siegel [**8**, **9**] initiated a line of research (see [**1**–**4**, **11**]) which generalized existing methods for treating $G(k)$. This typically led to upper bounds for $G_{\mathbb{K}}(k)$ of approximate strength $nB(k)$, where $B(k)$ was the best contemporary upper bound for $G(k)$. For example, Eda [**2**] gave an extension of Vinogradov's proof (see [**13**] or [**15**]) that $G(k) \leqslant (2+o(1))\,k \log k$. The present paper will eliminate the need for lengthy generalizations as such, by introducing a new and considerably shorter approach to the problem. Our main result is the following theorem.

THEOREM 1.1. *Let $\mathbb{K}$ be a number field of degree n over $\mathbb{Q}$. Then for large k one has*

$$G_{\mathbb{K}}(k) \leqslant nk\left(\log k + \log\log k + 6 + O\left(\frac{\log\log k}{\log k}\right)\right).$$

As with the bound $G(k) \leqslant (1+o(1))\,k \log k$ recently given by Wooley [**16**], Theorem 1.1 represents a halving improvement of current bounds when k is large. Our result may be compared with the bound given by Eda [**2**] for large k:

$$G_{\mathbb{K}}(k) \leqslant nk(2\log k + 3\log\log k + 4\log n + 2\log\log n) + 4k(\log n + 7) + 1. \tag{1.3}$$

Clearly, Theorem 1.1 improves (1.3) by a factor of $2 + 4\log n/\log k$. Our point of departure from Eda's work, as well as [**1**, **3**, **4**, **11**], is that we do not generalize to n dimensions the latest method found to be successful for $\mathbb{Q}$. At present this refers to the iterative method of Vaughan [**13**] with the efficient differencing procedure introduced by Wooley [**16**].

Received 7 March 1996; revised 3 October 1996.

1991 *Mathematics Subject Classification* 11P05 (primary), 12F05 (secondary).

J. London Math. Soc. (2) 59 (1999) 435–447

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

In view of Theorem 1.1 and Wooley's bound, one is tempted to prove that

$$G_{\mathbb{K}}(k) \leqslant nG(k). \tag{1.4}$$

Tatuzawa [**12**] proved that

$$G_{\mathbb{K}}(k) \leqslant 2n(G(k)+2k); \tag{1.5}$$

of course, in practice this yields $G_{\mathbb{K}}(k) \leqslant 2n(B(k)+2k)$, where $B(k)$ is currently Wooley's bound $k\log k(1+o(1))$. Inasmuch as this practical purpose is concerned, we will remove the leading factor of 2 in (1.5) by replacing the term $G(k)$ by a related function $G^{\alpha}(k)$. This new function will have two key features. First, $G^{\alpha}(k)$ is less than approximately $\frac{1}{2}B(k)$; see Lemma 1.3 below. Second, the inequality corresponding to (1.4), namely $G^{\alpha}_{\mathbb{K}}(k) < nG^{\alpha}(k)$, is not difficult to prove; see Lemma 4.2.

We need some terminology to define $G^{\alpha}(k)$. Given $0 < \alpha \leqslant 1$, we say that a set $\mathscr{S}$ of positive integers has exponential density α if

$$\operatorname{card}\{n \leqslant x : n \in \mathscr{S}\} > x^{\alpha-\epsilon}$$

for any $\epsilon > 0$ and all x sufficiently large (in terms of ϵ). We let $G^{\alpha}(k)$ be the least number s such that the set of positive integers which can be expressed as the sum of s kth powers of positive integers has exponential density α. The following result is somewhat more general than what is needed to prove Theorem 1.1.

THEOREM 1.2. *Suppose that* $k \geqslant 2$ *and* $1-(6-1/k)2^{-k} < \alpha \leqslant 1$. *Then for any number field* $\mathbb{K}$ *of degree* $n \geqslant 2$ *over* $\mathbb{Q}$, *we have*

$$G_{\mathbb{K}}(k) \leqslant 2n(G^{\alpha}(k)+2k).$$

Theorem 1.2 gives improved bounds for $G_{\mathbb{K}}(k)$ for each exponent $k \geqslant 4$ except $k = 5$, for n in finite ranges which may be calculated from results contained in [**1**, **12**], [**13**, Chapter 6], and Lemma 1.3 below. Since for small k there are numerous methods for obtaining still better results, we concern ourselves here only with the application to large k. We find that Theorem 1.1 improves known bounds by a factor of 2 for fields of degree less than approximately $2^k/k\log k$. This limitation arises from comparison of Theorem 1.1 with the uniform (in n) estimate

$$G_{\mathbb{K}}(k) \leqslant \max\{2^k+1, 8k^5\}$$

of Birch, Körner, and Ramanujam (see [**1**, **4**, **5**]).

Following custom, we let $G^{+}(k)$ be the least number s for which almost all positive integers are representable as a sum of s kth powers of non-negative integers. Clearly, one has $G^{1}(k) \leqslant G^{+}(k)$. It transpires that we can do no better for large k than to replace $G^{\alpha}(k)$ by $G^{+}(k)$ in our application of Theorem 1.2. In modern practice, if for large k a circle method argument yields $G(k) \leqslant B(k)$, then a related argument will produce

$$G^{+}(k) \leqslant \tfrac{1}{2}(B(k)+1).$$

For example, by altering the proof of the sharpest bound presently known for $G(k)$ for large k (see [**19**]), we prove the following main lemma.

LEMMA 1.3. *For large* k, *we have*

$$G^{+}(k) \leqslant \tfrac{1}{2}k\left(\log k + \log\log k + 2 + O\left(\frac{\log\log k}{\log k}\right)\right).$$

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

By appealing to bounds for $G^+(k)$, our approach offers a shortcut around the complicated details found in [**2**, **3**], which focus on Weyl sum estimates. Furthermore, it is clear that our strategy will improve bounds of the form $G_{\mathbb{K}}(k) \leqslant nB(k)$ whenever new circle method arguments give sharper bounds for $G(k)$. This leaves only the matter of reducing the dependency on n, which we will return to in a forthcoming paper. As a first step, we note that, unlike [**3**, Hauptsatz] and (1.3), Theorem 1.1 does not require extra terms involving $\log n$.

2. *Notation*

We let $\mathbb{K}$ be an algebraic number field of degree $n = r_1 + 2r_2$ over the rationals $\mathbb{Q}$, having r_1 real conjugate fields $\mathbb{K}^{(l)}$ $(1 \leqslant l \leqslant r_1)$ and r_2 pairs of complex conjugate fields $\mathbb{K}^{(m)}$ $(r_1+1 \leqslant m \leqslant r_1+r_2)$. The superscripts (l) and (m) will always refer to these ranges of indices. We say that $\mathbb{K}$ is totally real if $r_2 = 0$ and totally complex if $r_1 = 0$. Let $\mathscr{O} = \mathscr{O}_{\mathbb{K}}$ be the ring of integers of $\mathbb{K}$, let J_k be the ring generated by the kth powers of integers of $\mathbb{K}$, and denote the units of $\mathscr{O}$ by $\mathscr{U}$. Denote the different and absolute value of the discriminant of $\mathbb{K}$ by $\mathfrak{d}$ and D respectively. Then $\mathscr{N}(\mathfrak{d}) = D$, where $\mathscr{N}$ denotes the usual norm on ideals. Throughout this work, lowercase German (Fraktur) letters will denote integral ideals of $\mathbb{K}$. For γ in $\mathbb{K}$ we let $\gamma^{(i)}$ $(1 \leqslant i \leqslant n)$ be the conjugates of γ, where $\gamma^{(i)} \in \mathbb{K}^{(i)}$, and we denote the trace of γ by $\mathrm{Tr}(\gamma) = \sum_{1 \leqslant i \leqslant n} \gamma^{(i)}$ and the usual element-norm by $N(\gamma) = \prod_{1 \leqslant i \leqslant n} \gamma^{(i)}$. The conjugate height of γ is defined by $\|\gamma\| = \max_{1 \leqslant i \leqslant n} |\gamma^{(i)}|$. When $\lambda = \sum_{1 \leqslant j \leqslant n} t_j \gamma_j$ for $\gamma_j \in \mathbb{K}$ and real variables t_j $(1 \leqslant j \leqslant n)$, we define $\lambda^{(i)} = \sum_{1 \leqslant j \leqslant n} t_j \gamma_j^{(i)}$ for each i from 1 to n. We then extend the trace function so that $\mathrm{Tr}(\lambda) = \sum_{1 \leqslant i \leqslant n} \lambda^{(i)}$, and define $\|\lambda\|$ as above. For brevity we write $E(\gamma) = e^{2\pi i \mathrm{Tr}(\gamma)}$ and $E(\lambda) = e^{2\pi i \mathrm{Tr}(\lambda)}$, which is a generalization of Vinogradov's notation $e(x) = e^{2\pi i x}$.

Let $\{\omega_1, \omega_2, \ldots, \omega_n\}$ be an integral basis of $\mathscr{O}$, and denote by $\{\rho_1, \rho_2, \ldots, \rho_n\}$ a dual basis of $\mathfrak{d}^{-1}$ satisfying

$$\mathrm{Tr}(\rho_i \omega_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

As special cases of the linear form λ defined above, we let

$$\xi = \sum_{1 \leqslant i \leqslant n} t_i \rho_i, \quad \eta = \sum_{1 \leqslant i \leqslant n} y_i \omega_i.$$

When these forms appear in n-fold integrals, we use $d\mathbf{t}$ and $d\mathbf{y}$ to denote the n-fold differentials $dt_1\, dt_2 \ldots dt_n$ and $dy_1\, dy_2 \ldots dy_n$.

The symbols $\mathbb{R}$, $\mathbb{Z}$, $\mathbb{N}$ mean, as usual, the reals, rational integers, and positive rational integers. We let $\mathbb{T}^n$ be the unit n-cube in $\mathbb{R}^n$. If $\mathscr{F}$ is a finite set, we denote its cardinality by $\mathrm{card}\,\mathscr{F}$.

As usual, $\gamma \in \mathbb{K}$ is said to be totally positive or totally non-negative according to whether

$$\gamma^{(l)} > 0 \quad (1 \leqslant l \leqslant n) \quad \text{or} \quad \gamma^{(l)} \geqslant 0 \quad (1 \leqslant l \leqslant n),$$

that is, if all of its real conjugates are positive or non-negative respectively. By default, the term positive when used alone will always refer to a positive real number. Denote by $\mathbb{K}^+$ and $\mathscr{O}_{\mathbb{K}}^+$ the sets of totally non-negative elements and non-negative integers

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

(respectively) of $\mathbb{K}$. Note that if $\mathbb{K}$ is totally complex then $\mathbb{K} = \mathbb{K}^+$ (vacuously). For a positive parameter T we define the sets $\mathscr{S}(T)$, where the symbol $\mathscr{S}$ is $\mathbb{K}$, $\mathbb{K}^+$, $\mathscr{O}$, or $\mathscr{O}^+$, by $\mathscr{S}(T) = \{\gamma \in \mathscr{S} : \|\gamma\| \leqslant T\}$. Furthermore we let $P(T)$ denote the set of all $\mathbf{y} \in \mathbb{R}^n$ such that $\eta = \eta(\mathbf{y})$ (defined above) satisfies $\|\eta\| \leqslant T$ and $0 \leqslant \eta^{(l)} \leqslant T$ $(1 \leqslant l \leqslant n)$. Hence $\mathbb{K}(T)$ and $\mathscr{O}(T)$ are subsets of $\mathbb{K}$, whereas $P(T)$ is a subset of $\mathbb{R}^n$.

For a complex number s, the real part of s is denoted by $\Re s$. We will make frequent use of the $\ll$ and $\gg$ notations of Vinogradov, as well as the O and o symbols of Landau, wherein all implied constants may depend on $\mathbb{K}$, k, and s, unless stated otherwise. The dependence on $\mathbb{K}$ will be made explicit whenever it affects the bounds we achieve for $G_{\mathbb{K}}(k)$. We use ϵ to denote a small positive number whose value need not be the same in consecutive appearances.

3. *Outline of Siegel's circle method*

In 1921, proceeding along the lines of Hardy–Littlewood's work and implementing Hecke's theta-series, Siegel [**6**, **7**] proved that $G_{\mathbb{K}}(2) \leqslant 5$ for all $\mathbb{K}$. He did not find a suitable number field generalization of the Farey dissection, allowing a treatment of $G_{\mathbb{K}}(k)$ for all $k \geqslant 2$, until the 1940s. We now outline his later treatment (see [**8**, **9**]), utilizing improvements of Stemmler [**10**] and Tatuzawa [**11**]. We refer the reader to [**16**] for proofs and their original references. The following lemma generalizes a simple orthogonality relation for $\mathbb{Z}$ which is used to detect solutions of Diophantine equations.

LEMMA 3.1. *Let γ be an integer of $\mathbb{K}$. Then*

$$\int_{\mathbb{T}^n} E(\gamma\xi)\, d\mathbf{t} = \begin{cases} 1 & \text{if } \gamma = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Proof. See [**16**, Lemma 5.2]. □

Fixing a totally non-negative integer v, we let $R_{s,k,T}(v)$ be the number of solutions of the Diophantine equation

$$v = \lambda_1^k + \ldots + \lambda_s^k, \quad \lambda_i \in \mathscr{O}^+(T)\ (1 \leqslant i \leqslant s).$$

Then if we define the generating function

$$F(\xi, T) = \sum_{\lambda \in \mathscr{O}^+(T)} E(\lambda^k \xi),$$

it follows from Lemma 3.1 that

$$R_{s,k,T}(v) = \int_{\mathbb{T}^n} F(\xi, T)^s E(-v\xi)\, d\mathbf{t}.$$

By using Minkowski's theorem in place of the pigeonhole principle, Siegel extended Dirichlet's theorem on rational approximation to number fields, and

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

thereby defined the *basic domain*, generalizing the major arcs, as consisting of points in $\mathbb{T}^n$ which are well-approximable in a certain sense by elements of $\mathbb{K}$. The next lemma makes this precise. In what follows, we let h and t denote real numbers satisfying

$$t > 1, \quad h > 2D^{1/n}t.$$

LEMMA 3.2. *Fix a number field $\mathbb{K}$ and suppose that t and h are chosen as above. Then corresponding to any ξ, there exist $\alpha \in \mathcal{O}$ and $\beta \in \mathfrak{d}^{-1}$ such that*

$$\|\alpha\xi - \beta\| < \frac{1}{h}, \quad t < \|\alpha\| \leqslant h,$$

$$\max(h\,|\,\alpha^{(i)}\xi^{(i)} - \beta^{(i)}|, |\alpha^{(i)}|) \geqslant \frac{1}{\sqrt{D}} \quad (1 \leqslant i \leqslant n)$$

and

$$\mathcal{N}((\alpha, \beta\mathfrak{d})) \leqslant \sqrt{D}.$$

Proof. See [**S4**, Lemma 6], or [**16**, Theorem 3.1]. □

For γ in $\mathbb{K}$, we can find unique integral ideals $\mathfrak{a}$ and $\mathfrak{b}$ with $(\mathfrak{a}, \mathfrak{b}) = 1$ such that $\gamma\mathfrak{d} = \mathfrak{b}/\mathfrak{a}$. This relation is denoted by $\gamma \to \mathfrak{a}_\gamma$, and we recall $\mathfrak{a} = \mathfrak{a}_\gamma$ the denominator of γ.

Recall that $\mathbb{T}$ denotes $[0, 1]$. We define $\Gamma(t)$ to be the set of γ which, when expressed via the dual basis $\{\rho_i\}_{i=1}^n$ as $\gamma = \sum_{i=1}^n u_i\rho_i$, satisfy

$$(u_1, \ldots, u_n) \in (\mathbb{T} \cap \mathbb{Q})^n \quad \text{and} \quad \mathcal{N}(\mathfrak{a}_\gamma) \leqslant t^n.$$

For $\gamma \in \Gamma(t)$, we write $\xi = \sum_{i=1}^n t_i\rho_i$ and define B_γ to be the set of all $\mathbf{t} \in \mathbb{T}^n$ which satisfy, for some $\gamma_0 \equiv \gamma \pmod{\mathfrak{d}^{-1}}$,

$$\prod_{i=1}^n \max\left(h|\xi^{(i)} - \gamma_0^{(i)}|, \frac{1}{t}\right) \leqslant \frac{1}{\mathcal{N}(\mathfrak{a})}.$$

Siegel defined the basic domain corresponding to t and h by

$$B = B(t, h) = \bigcup_{\gamma \in \Gamma(t)} B_\gamma.$$

Analogously to the minor arcs we have the complement

$$S = S(t, h) = \mathbb{T}^n \backslash B,$$

which is called the *supplementary domain*. Taken together, the basic domain and the supplementary domain form a generalized Farey dissection of $\mathbb{T}^n$ into B and S with respect to (t, h).

Observe that B_γ is empty when $\mathcal{N}\mathfrak{a}_\gamma > t^n$, and that, in contrast to the rational case, B_γ is typically not an n-dimensional interval. When $n = 1$ here, B_γ is the major arc usually denoted by $\mathcal{M}(a, q)$, since the condition defining $\mathcal{M}(a, q)$ may be written as

$$\max\left(h\left|\alpha - \frac{a}{q}\right|, \frac{1}{t}\right) \leqslant \frac{1}{q}.$$

As with the major arcs, the basic domains are disjoint.

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

LEMMA 3.3. *Suppose that γ_1 and γ_2 are distinct elements of $\Gamma(t)$. Then*

$$B_{\gamma_1} \cap B_{\gamma_2} = \varnothing.$$

Proof. See [**16**, Lemma 5.1]. □

In Siegel's method there appear generalizations, defined below, of the singular integral $J(n)$ and singular series $\mathfrak{S}(n)$. For $\gamma \in \Gamma(t)$ with $\gamma \to \mathfrak{a}$, we set

$$\zeta = \xi - \gamma,$$
$$\mathscr{G}(\gamma) = N(\mathfrak{a})^{-1} \sum_{\lambda (\mathrm{mod}\, \mathfrak{a})} E(\lambda^k \gamma),$$
$$I(\zeta, T) = \int_{P(T)} E(\eta^k \zeta)\, d\mathbf{y},$$

where (as before) $\eta = \sum_{i=1}^n y_i \omega_i$. Then for $\mu \in \mathbb{K}$ and $\nu \in \mathscr{O}$ we define the singular integral $J(\mu)$ and singular series $\mathfrak{S}(\nu)$ by

$$\begin{aligned} J(\mu) &= \int_{\mathbb{R}^n} I(\zeta)^s E(-\mu\zeta)\, d\mathbf{t}, \\ \mathfrak{S}(\nu) &= \sum_{\gamma (\mathrm{mod}\, \mathfrak{d}^{-1})} \mathscr{G}(\gamma)^s E(-\nu\gamma). \end{aligned} \tag{3.1}$$

Since, by definition, $\mathfrak{d}^{-1} = \{\rho : E(\rho\alpha) = 1 \text{ for every } \alpha \in \mathscr{O}\}$, we have $E(\lambda^k\gamma) = E(\mu^k\gamma)$ provided that $\gamma \to \mathfrak{a}$ and $\lambda, \mu \in \mathscr{O}$ with $\lambda \equiv \mu \pmod{\mathfrak{a}}$. Hence the sums defining $\mathscr{G}(\gamma)$ and $\mathfrak{S}(\nu)$ are independent of the choice of a residue system.

LEMMA 3.4. *Assume that $s \geqslant 4kn$ and $\nu \in J_k$, where $k \geqslant 3$. Then there exist positive constants c_1, c_2 depending on s, k, $\mathbb{K}$ such that $c_1 < \mathfrak{S}(\nu) < c_2$.*

Proof. See [**16**, Theorem 7.1]. □

For the singular integral, we have the following lemma.

LEMMA 3.5. *If $s > k$ and $\mu = \nu T^{-k} \in \mathbb{K}^+(1)$, then the singular integral $J(\mu)$ is absolutely convergent, and*

$$J(\mu) = D^{(1-s)/2} k^{-sn} \Gamma\left(\frac{1}{k}\right)^{sr_1} \Gamma\left(\frac{s}{k}\right)^{-r_1} \prod_{(l)} (\mu^{(l)})^{s/k-1} \prod_{(m)} H(\mu^{(m)})$$

where

$$H(\mu^{(m)}) = \int_{V_m} \prod_{i=1}^{s} w_i^{1/k-1}\, dw_1 \ldots dw_{s-1}\, d\phi_1 \ldots d\phi_{s-1}$$

with

$$w_s = |\mu^{(m)} - w_1^{1/2} e^{i\phi_1} \ldots - w_{s-1}^{1/2} e^{i\phi_{s-1}}|^2$$

and V_m is the region determined by

$$0 \leqslant w_i \leqslant 1 \quad (1 \leqslant i \leqslant s), \quad -\pi \leqslant \phi_j \leqslant \pi \quad (1 \leqslant j \leqslant s-1).$$

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

Proof. See [**16**, Theorem 6.2]. □

Finally, for $k \geqslant 2$ we set

$$a = \frac{1}{4} + \frac{1}{4kn}, \quad t' = T^{1-a}, \quad h' = T^{k-1+a}, \tag{3.2}$$

and form the generalized Farey dissection of $\mathbb{T}^n$ into B' and S' with respect to (t', h'). Then the above results may be combined into the following theorem, which expresses the basic domain integral arising from the parameters (3.2) asymptotically in terms of arithmetic functions.

THEOREM 3.6. *Let ν be a totally positive integer of $\mathbb{K}$, and set $\mu = \nu T^{-k}$. Then if $s \geqslant 4kn$, we have*

$$\int_{B'} F(\xi, T)^s E(-\nu\xi)\, d\mathbf{t} = \mathfrak{S}(\nu)\, J(\mu)\, T^{n(s-k)} + O(T^{n(s-k)-a}).$$

Proof. See [**16**, Theorem 5.1]. □

The deepest part of Siegel's method involves showing that, as $T \to \infty$,

$$\int_S F(\xi, T)^s E(-\nu\xi)\, d\mathbf{t} = o\left(\int_B F(\xi, T)^s E(-\nu\xi)\, d\mathbf{t} \right)$$

provided s is sufficiently large in terms of k and n. This requires a generalized Weyl sum estimate; the following one will be sufficient for our purposes.

LEMMA 3.7. *Let*

$$f(\lambda) = \alpha_k \lambda^k + \ldots + \alpha_1 \lambda$$

be a polynomial of degree k with coefficients in $\mathcal{O}$ and with constant term zero. Given $T > 1$, suppose that h is a number satisfying $T^{k-1} \ll h \ll T^k$. Then for fixed ξ, and given any $\epsilon > 0$, one has

$$\left| \sum_{\lambda \in \mathcal{O}^+(T)} E(f(\lambda)\xi) \right| \ll T^{n+\epsilon} \left(\frac{1}{\|\alpha\|} + \frac{1}{T} + \frac{h}{T^k} \right)^{2^{1-k}}$$

where α denotes an integer satisfying the conditions in Lemma 3.2 *corresponding to $\alpha_k \xi$, and the implicit constants depend on k, $\mathbb{K}$, and ϵ.*

Proof. See [**16**, Theorem 3.2]. □

COROLLARY 3.8. *Let θ be a number satisfying $0 < \theta < (1-a)\, 2^{1-k}$, where a is given by* (3.2). *Then for $\mathbf{t} \in S'$, we have*

$$|F(\xi, T)| \ll T^{n-\theta}$$

where as usual $\xi = \sum_{i=1}^n t_i \rho_i$, and the implicit constant depends only on θ.

Proof. See [**16**, Lemma 5.10]. □

4. *Proof of Theorem* 1.2

Tatuzawa's argument relied on a lemma which we restate here for convenience.

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

LEMMA 4.1. *Suppose that $k \in \mathbb{N}$ and $\mathbb{K}$ is a number field of degree n over $\mathbb{Q}$. Let T be a sufficiently large positive number, and denote by $W_s(T)$ the set of integers of $\mathbb{K}$ which can be represented in the form*

$$\lambda_1^k + \ldots + \lambda_s^k$$

with

$$\lambda_i \in \mathcal{O}^+(T) \quad (1 \leqslant i \leqslant s).$$

Then if $s_1 = nG(k)$, we have

$$\operatorname{card} W_{s_1}(T) \gg T^{nk}$$

where the implicit constant depends only on k and $\mathbb{K}$.

Proof. See [**16**, Lemma 7.6]. □

We require a modification of this lemma invoking $G^{\alpha}(k)$ in place of $G(k)$.

LEMMA 4.2. *Adopt the same hypotheses as in the statement of Lemma* 4.1. *In analogy with $G^{\alpha}(k)$, define $G^{\alpha}_{\mathbb{K}}(k)$ to be the least s such that* $\operatorname{card} W_s(T) \gg T^{nk\alpha-e}$ *for any $\epsilon > 0$. Then for any $0 < \alpha \leqslant 1$ one has*

$$G^{\alpha}_{\mathbb{K}}(k) \leqslant nG^{\alpha}(k).$$

Proof. Select $\theta \in \mathcal{O}_{\mathbb{K}}$ such that $1, \theta, \ldots, \theta^{n-1}$ are linearly independent over $\mathbb{Q}$, and let ζ_k be a kth root of unity other than 1. Choose a rational integer z large enough to ensure that $\theta + z$ is totally non-negative and that the conjugates satisfy

$$\theta^{(i)} + z \neq (\theta^{(j)} + z)\zeta_k \quad (1 \leqslant i < j \leqslant n)$$

for every kth root of unity ζ_k. It follows that the Vandermonde determinant

$$\det \begin{vmatrix} 1 & (\theta^{(1)}+z)^k & \ldots & (\theta^{(1)}+z)^{(n-1)k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (\theta^{(n)}+z)^k & \ldots & (\theta^{(n)}+z)^{(n-1)k} \end{vmatrix} = \prod_{1 \leqslant i \leqslant n} ((\theta^{(i)}+z)^k - (\theta^{(j)}+z)^k)$$

is non-zero, so we may assume that the totally positive integers

$$1, (\theta+z)^k, (\theta+z)^{2k}, \ldots, (\theta+z)^{(n-1)k}$$

are linearly independent over the rationals. Consider integers of the form

$$\gamma = x_0 + x_1(\theta+z)^k + \ldots + x_{n-1}(\theta+z)^{k(n-1)}, \tag{4.1}$$

where the coefficients x_j are natural numbers and each term satisfies

$$x_j(\theta+z)^{jk} \in \mathcal{O}^+(T^k) \quad (0 \leqslant j \leqslant n-1).$$

If we let

$$c_j = \|(\theta+z)^{kj}\|^{-1},$$

then we have $x_j \leqslant c_j T^k$ for each $0 \leqslant j \leqslant n-1$. Clearly, each coefficient x_j may be chosen as a sum of $G^{\alpha}(k)$ kth powers of natural numbers in at least $(c_j T^k)^{\alpha-\epsilon}$ ways, for T sufficiently large and any $\epsilon > 0$. Since all such numbers of the form (4.1) are distinct by linear independence, we can construct more than $c_0 c_1 \ldots c_{n-1}(T^{k\alpha-\epsilon})^n$ integers $\gamma \in W_{s_1}(T)$ in this manner, where $s_1 = nG^{\alpha}(k)$. The lemma follows. □

The proof of Theorem 1.2 requires the following well-known consequence of Dirichlet's theorem on units in number fields.

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

LEMMA 4.3. *Recall that $r = r_1 + r_2 - 1$, and let $b_1, \dots, b_{r+1}$ be a set of real numbers satisfying*

$$\sum_l b_l + 2\sum_m b_m = 0.$$

Then there exists a totally non-negative unit η of $\mathbb{K}$ such that

$$c^{-1}e^{b_i} \leqslant |\eta^{(i)}| \leqslant ce^{b_i} \quad (1 \leqslant i \leqslant r+1),$$

where c is a positive constant depending only on $\mathbb{K}$.

Proof. See [**16**, Lemma 3.1]. □

Proof of Theorem 1.2. The proof, which utilizes Siegel's generalization of the Hardy–Littlewood method, is similar to that of Tatuzawa's bound (1.4); see [**16**, Theorem 7.3]. By Siegel's result $G_{\mathbb{K}}(2) \leqslant 5$ (see [**7**]), we may suppose that $k \geqslant 3$. Set $s = 4kn$. Given a non-zero element $v \in J_k \cap \mathcal{O}^+$, we substitute

$$b_i = k^{-1}(\log N(v)^{1/n} + \log |v^{(i)}|^{-1}) \quad (1 \leqslant i \leqslant n)$$

in Lemma 4.3, and obtain a unit $\eta \in \mathcal{O}^+$ satisfying $e^{b_i} \ll |\eta^{(i)}| \ll e^{b_i}$ $(1 \leqslant i \leqslant n)$. Then $\chi = v\eta^k$ satisfies

$$c_1 M \leqslant \chi^{(l)} \leqslant c_2 M \quad (1 \leqslant l \leqslant r_1),$$
$$c_1 M \leqslant |\chi^{(m)}| \leqslant c_2 M \quad (1 \leqslant m \leqslant r_1 + r_2)$$

where $M = N(v)^{1/n}$ and c_1 and c_2 are constants depending only on k and $\mathbb{K}$. Let $s_1 = nG^*(k)$, and take $\sigma_1, \sigma_2 \in W_{s_1}(T_1)$, where

$$T_1 = \left(\frac{c_1 M}{4s_1}\right)^{1/k}.$$

Also set

$$T = ((c_2 + \tfrac{1}{4}c_1) M)^{1/k}.$$

Now define the generating function

$$V(\xi, T) = \sum_{\sigma \in W_{s_1}(T)} E(\sigma\xi),$$

and let $V(\xi)$ denote $V(\xi, T_1)$. Recall also that

$$F(\xi, T) = \sum_{\lambda \in \mathcal{O}^+(T)} E(\lambda^k \xi),$$

which we shall abbreviate to $F(\xi)$. Let $\tau = \chi - \sigma_1 - \sigma_2$. Taking the Farey dissection of $\mathbb{T}^n$ into B' and S' with respect to the parameters t', h' in (3.2), we have, from Theorem 3.6,

$$\int_{B'} F(\xi)^s V(\xi)^2 E(-\chi\xi)\, d\mathbf{t} = \sum_\tau \int_{B'} F(\xi)^s E(-\tau\xi)\, d\mathbf{t}$$
$$= \sum_\tau \mathfrak{S}(v) J(\mu) T^{n(s-k)}(1 + o(1)),$$

where $\mu = \tau T^{-k}$ and $\mathfrak{S}(v)$ and $J(\mu)$ are defined by (3.1). It follows from the choice of T above that $\mu \in \mathbb{K}^+(1)$, so Lemma 3.5 shows that

$$J(\mu) > c > 0$$

for some constant $c = c(k, \mathbb{K})$. Similarly we have, from Lemma 3.4,

$$\mathfrak{S}(v) > c_1(k, \mathbb{K}) > 0.$$

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

Therefore the real part of the basic domain integral satisfies

$$\Re \int_{B'} F(\xi)^s V(\xi)^2 E(-\chi\xi)\, d\mathbf{t} \gg T^{n(s-k)} (\operatorname{card} W_{s_1}(T_1))^2,$$

which by Lemma 4.2 is

$$\gg T^{n(s-k(1-\alpha))-\epsilon} \cdot \operatorname{card} W_{s_1}(T_1) \tag{4.2}$$

for any $\epsilon > 0$. From Lemma 3.1 we have

$$\int_{\mathbb{T}^n} |V(\xi)|^2 d\mathbf{t} = \int_{\mathbb{T}^n} \sum_{\sigma_1, \sigma_2 \in W_{s_1}(T_1)} E((\sigma_1 - \sigma_2)\xi)\, d\mathbf{t}$$
$$= \operatorname{card} W_{s_1}(T_1).$$

Since $n \geqslant 2$, the parameter a in (3.2) satisfies $a \leqslant 1/4 + 1/(8k)$, and therefore we may apply Corollary 3.8 with $\theta = (6 - 1/k)\, 2^{-k-2} - \epsilon'$ with a small ϵ' to obtain

$$\int_{S'} F(\xi)^s V(\xi)^2 E(-\chi\xi)\, d\mathbf{t} \ll T^{s(n-\theta)} \cdot \operatorname{card} W_{s_1}(T_1). \tag{4.3}$$

From the hypotheses $s = 4nk$ and $1 - \alpha < (6 - 1/k)\, 2^{-k} = 4\theta + \varepsilon'$, we see that

$$s(n-\theta) < n(s - k(1-\alpha)) - \epsilon$$

for ϵ sufficiently small. Hence by comparing (4.2) and (4.3), it follows that for T and T_i large enough,

$$\int_{\mathbb{T}^n} F(\xi)^s V(\xi)^2 E(-\chi\xi)\, d\mathbf{t} > 0.$$

By Lemma 3.1, we conclude that $\chi = v\eta^k$ can be expressed as

$$\chi = \lambda_1^k + \ldots + \lambda_s^k + \tau_1^k + \ldots + \tau_{2s_1}^k,$$

where

$$\lambda_i \in \mathcal{O}^+(T) \quad (1 \leqslant i \leqslant s) \quad \text{and} \quad \tau_j \in \mathcal{O}^+(T_1) \quad (1 \leqslant j \leqslant 2s_1).$$

Since η is a totally non-negative unit, $s = 4kn$, and $s_1 = nG^{\alpha}(k)$, the proof of Theorem 1.2 is complete. □

5. *Proof of Theorem* 1.1

Theorem 1.1 follows immediately from Theorem 1.2 and Lemma 1.3. We now prove the latter by a well-known argument as given in [**17**, Section 8], but utilizing the more convenient and sharper estimates contained in [**19**].

Proof of Lemma 1.3. Let N be a large natural number, and set $P = N^{1/k}$. For a positive parameter R with $R \leqslant P$, let $\mathscr{A}(P, R)$ be the set of *R-smooth* positive integers up to P, that is,

$$\mathscr{A}(P, R) = \{1 \leqslant n \leqslant P : p \text{ prime}, p \mid n \Rightarrow p \leqslant R\}.$$

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

Letting η be a small positive number, we set

$$f(\alpha; P, \eta) = \sum_{x \in \mathscr{A}(P, P^{\eta})} e(\alpha x^k),$$

$$g(\alpha; P) = \sum_{1 \leqslant x \leqslant P} e(\alpha x^k),$$

and abbreviate these to $f(\alpha)$ and $g(\alpha)$ respectively. Let v, w be natural numbers with $v \geqslant k-1$, and set $s = v+w$. We denote by $R(n)$ the number of representations of the positive integer n in the form

$$n = x_1^k + x_2^k + y_1^k + \ldots + y_s^k,$$

subject to

$$1 \leqslant x_1, x_2 \leqslant P \quad \text{and} \quad y_i \in \mathscr{A}(P, P^{\eta}) \quad (1 \leqslant i \leqslant s).$$

Then

$$R(n) = \int_0^1 g(\alpha)^2 f(\alpha)^s e(-\alpha n)\, d\alpha.$$

Define $\mathfrak{m}$ to be the set of real numbers $\alpha \in [0,1)$ such that, whenever $a \in \mathbb{Z}$, $q \in \mathbb{N}$, $(a,q) = 1$ and $|\alpha - a/q| \leqslant q^{-1} P^{1-k}$, one has $q > P$. Then $\mathfrak{M} = [0,1) \backslash \mathfrak{m}$.

Next, we define

$$R_{\mathfrak{m}}(n) = \int_0^1 F(\alpha)\, e(-\alpha n)\, d\alpha$$

to be the nth Fourier coefficient of the function

$$F(\alpha) = \begin{cases} g(\alpha)^2 f(\alpha)^s & \text{if } \alpha \in \mathfrak{m} \\ 0 & \text{otherwise.} \end{cases}$$

Also, let

$$R_{\mathfrak{M}}(n) = \int_{\mathfrak{M}} g(\alpha)^2 f(\alpha)^s e(-\alpha n)\, d\alpha.$$

By Bessel's inequality, for each natural number N we have

$$\begin{aligned} \sum_{n \leqslant N} |R(n) - R_{\mathfrak{M}}(n)|^2 &= \sum_{n \leqslant N} |R_{\mathfrak{m}}(n)|^2 \\ &\leqslant \int_0^1 |F(\alpha)|^2\, d\alpha \\ &= \int_{\mathfrak{m}} |g(\alpha)^4 f(\alpha)^{2s}|\, d\alpha \\ &\leqslant P^2 (\sup_{\alpha \in \mathfrak{m}} |f(\alpha)|)^{2w} \int_0^1 |g(\alpha)^2 f(\alpha)^{2v}|\, d\alpha. \end{aligned}$$

By the corollary to [**18**, Theorem 2.1] and the remark at the end of [**17**, Section 3], one has

$$\int_0^1 |g(\alpha)^2 f(\alpha)^{2v}|\, d\alpha \ll P^{2v-k+\Delta_{v+1}+\epsilon},$$

where

$$\Delta_{v+1} = k e^{1-2v/k}.$$

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

Also, by Corollary 2 of [**19**, Theorem 4.2], we have

$$\sup_{\alpha\in\mathfrak{m}}|f(\alpha)| \ll P^{1-\sigma(k)+\epsilon},$$

where

$$\sigma(k)^{-1} = k(\log k + O(\log\log k)). \tag{5.1}$$

Then by taking

$$w = \left\lfloor \frac{\Delta_{v+1}}{2\sigma(k)} \right\rfloor + 1,$$

we will have

$$\int_{\mathfrak{m}} |g(\alpha)^4 f(\alpha)^{2(v+w)}|\, d\alpha \ll P^{2s+4-k-\delta}$$

for some $\delta > 0$. Meanwhile, the pruning argument used in [**14**, Section 5] applies to the major arcs $\mathfrak{M}$, due to the presence of the factor $g(\alpha)^2$, and for $N/2 < n \leqslant N$ (say) one obtains

$$\int_{\mathfrak{M}} g(\alpha)^2 f(\alpha)^s e(-\alpha n)\, d\alpha \gg P^{s+2-k}, \tag{5.2}$$

provided that $s+2 \geqslant 4k$. This last requirement ensures the positivity of the singular series, and may be freely assumed since we later choose $v \gg k\log k$. It follows that

$$\begin{aligned}\sum_{N/2<n\leqslant N} \left|\frac{R(n)-R_{\mathfrak{M}}(n)}{R_{\mathfrak{M}}(n)}\right|^2 &\ll \frac{N^{(2s+4-k-\delta)/k}}{(N^{(s+2-k)/k})^2} \\ &= N^{1-\delta/k}.\end{aligned}$$

Letting $N\to\infty$, we conclude that almost no integers n in $[1,N]$ have $R(n)=0$. Hence for large k we have shown that

$$G^+(k) \leqslant \max\left\{v+3+\left\lfloor \frac{\Delta_{v+1}}{2\sigma(k)}\right\rfloor, 4k\right\}.$$

Choosing

$$v = [\tfrac{1}{2}k(\log k + \log\log k + 1)],$$

one sees that $\Delta_{v+1} \leqslant 1/\log k$. The desired explicit bound now follows from (5.1). □

Acknowledgements. I extend my deepest gratitude to Trevor Wooley for his guidance during and beyond the completion of my PhD thesis (University of Michigan), of which an earlier version of this paper forms a chapter. I also wish to thank him for financial support through his grant from the David and Lucile Packard Foundation. The submitted version of this paper was completed while the author was supported by NSF grant DMS 9304580 while at the Institute for Advanced Study, whose hospitality is much appreciated. Finally, I thank the referee for several helpful comments.

References

1. B. J. Birch, 'Waring's problem in algebraic number fields', *Proc. Cambridge Philos. Soc.* 57 (1961) 449–459.
2. Y. Eda, 'On the Waring problem in an algebraic number field', *Seminar on Modern Methods in Number Theory* (Tokyo, 1971).
3. O. Körner, 'Über die Waringsche Problem in algebraischen Zahlkörper', *Math. Ann.* 144 (1961) 224–238.

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press

4. O. Körner, 'Über Mittlewerte trigometrischer Summen und ihre Anwendung in algebraischen Zahlkörpern', *Math. Ann.* 147 (1962) 205–239.
5. C. P. Ramanujam, 'Sums of *m*th powers in *p*-adic rings', *Mathematika* 10 (1963) 137–146.
6. C. L. Siegel, 'Darstellung total positiver Zahlen durch Quadrate', *Math. Z.* 10 (1921) 246–275; or see *Ges. Abh. I* (Springer, 1966) 47–76.
7. C. L. Siegel, 'Additive Theorie der Zahlkörper II', *Math. Ann.* 88 (1923) 184–210; or see *Ges. Abh. I* (Springer, 1966) 180–206.
8. C. L. Siegel, 'Generalization of Waring's problem to algebraic number fields', *Amer. J. Math.* 66 (1944) 122–136; or see *Ges. Abh. II* (Springer, 1966) 406–420.
9. C. L. Siegel, 'Sums of *m*th powers of algebraic integers', *Ann. Math.* 46 (1945) 313–339; or see *Ges. Abh. III* (Springer, 1966) 12–38.
10. R. M. Stemmler, 'The easier Waring problem in algebraic number fields', *Acta Arith.* 6 (1961) 447–468.
11. T. Tatuzawa, 'On the Waring problem in an algebraic number field', *J. Math. Soc. Japan* 10 (1958) 322–341.
12. T. Tatuzawa, 'On Waring's problem in algebraic number fields', *Acta Arith.* 24 (1973) 37–60.
13. R. C. Vaughan, *The Hardy–Littlewood method* (Cambridge University Press, Cambridge, 1981).
14. R. C. Vaughan, 'A new iterative method in Waring's problem', *Acta Math.* 162 (1989) 1–71.
15. I. M. Vinogradov, 'On the question of an upper bound for $G(n)$' *Izv. Akad. Nauk.* 23 (1959) 637–642 (in Russian).
16. Y. Wang, *Diophantine equations and inequalities in algebraic number fields* (Springer, Berlin, 1991).
17. T. D. Wooley, 'Large improvements in Waring's problem', *Ann. Math.* 135 (1992) 131–164.
18. T. D. Wooley, 'The application of a new mean value theorem to the fractional parts of polynomials', *Acta Arith.* 65 (1993) 163–179.
19. T. D. Wooley, 'New estimates for smooth Weyl sums', *J. London Math. Soc.* (2) 51 (1995) 1–13.

Department of Mathematics and Computer Science
Kent State University
Kent
OH 44242
USA

https://doi.org/10.1112/S0024610799007309 Published online by Cambridge University Press