ILLINOIS JOURNAL OF MATHEMATICS
Volume 27, Number 1, Spring 1983

DIOPHANTINE APPROXIMATION AND NORM DISTRIBUTION IN GALOIS ORBITS

BY

C. J. BUSHNELL

This paper investigates another facet of a problem, introduced in [1] and [2], concerning relations (or lack of them) between arithmetic properties and Galois-module properties of algebraic numbers. We let K be an algebraic number field, normal (and of finite degree) over the field $\mathbf{Q}$ of rational numbers, and we write $\Gamma = \mathrm{Gal}(K/\mathbf{Q})$. The natural action $\gamma: b \mapsto b^{\gamma}$, $\gamma \in \Gamma$, $b \in K$, of the Galois group on K extends to a module structure over the rational group algebra $\mathbf{Q}\Gamma$. Explicitly, if $a \in K$, and if

$$x = \sum_{\gamma \in \Gamma} x_{\gamma}\gamma, \quad x_{\gamma} \in \mathbf{Q},$$

is a typical element of $\mathbf{Q}\Gamma$, the module action is given by the formula

$$a \cdot x = \sum_{\gamma \in \Gamma} a^{\gamma} x_{\gamma}.$$

One knows (Hilbert's Normal Basis Theorem) that K is a free $\mathbf{Q}\Gamma$-module of rank one. That is, there exists $a \in K$ such that $K = a \cdot \mathbf{Q}\Gamma$, or, equivalently, the conjugates a^{γ}, $\gamma \in \Gamma$, of a are linearly independent over $\mathbf{Q}$.

This module structure naturally leads to others. Most notably, the ring $\mathfrak{o}_K$ of algebraic integers in K is a module over the integral group ring $\mathbf{Z}\Gamma$. If every prime of $\mathbf{Q}$ is at most tamely ramified in K, $\mathfrak{o}_K$ is "usually" a free $\mathbf{Z}\Gamma$-module: there exists $a \in K$ such that $\mathfrak{o}_K = a \cdot \mathbf{Z}\Gamma$. (This holds if, for example, Γ has no irreducible symplectic characters. See [4] for a complete account.) It is the arithmetic properties of these elements a with $a \cdot \mathbf{Z}\Gamma = \mathfrak{o}_K$ which primarily interest us. Here we are concerned with their norms.

This is better considered in a more general context. We fix an element $a \in K$ such that $a \cdot \mathbf{Q}\Gamma = K$. (These elements are, in some geometrical sense, typical.) The linear isomorphism $\mathbf{Q}\Gamma \simeq K$ given by $x \mapsto a \cdot x$, $x \in \mathbf{Q}\Gamma$, enables us to transfer arithmetical functions from K to $\mathbf{Q}\Gamma$. In particular, we write

$$\nu_a(x) = N_{K/\mathbf{Q}}(a \cdot x), \quad x \in \mathbf{Q}\Gamma,$$

Received March 27, 1981.

where $N_{K/\mathbf{Q}}$ denotes the field norm. We consider the distribution of the numbers $\nu_a(x)$ as x ranges over *the group* $\mathbf{Z}\Gamma^\times$ *of invertible elements of the integral group ring* $\mathbf{Z}\Gamma$. Observe that in the case $\mathfrak{o}_K = a \cdot \mathbf{Z}\Gamma$ above, we have $\mathfrak{o}_K = a' \cdot \mathbf{Z}\Gamma$ if and only if there exists $x \in \mathbf{Z}\Gamma^\times$ such that $a' = a \cdot x$. In general, if we have $a, b \in K$ with $a \cdot \mathbf{Q}\Gamma = b \cdot \mathbf{Q}\Gamma = K$, then the equation $a \cdot x = b$ has a solution $x \in \mathbf{Z}\Gamma^\times$ if and only if the free $\mathbf{Z}\Gamma$-lattices $a \cdot \mathbf{Z}\Gamma$, $b \cdot \mathbf{Z}\Gamma$, in K are equal. Thus we are concerned with an arithmetical property (norm distribution) of elements specified by a Galois-module property (generating a given $\mathbf{Z}\Gamma$-lattice).

We arrange the investigation around the question of whether or not the following statement holds:

Given $M > 0$, and $a \in K$ such that $a \cdot \mathbf{Q}\Gamma = K$, the inequality $|\nu_a(x)| < M$ has only finitely many solutions $x \in \mathbf{Z}\Gamma^\times$.

We give some quantitative results, for certain abelian groups Γ, which compare $\nu_a(x)$ with the height of the algebraic number $a \cdot x$, and imply a positive answer to this question for these Γ. Our principal tool is a theorem of W. M. Schmidt (quoted in full in §1) which describes completely the behaviour of products of linear forms with algebraic coefficients. This enables us to prove a startlingly precise result when Γ has prime order. The method is not so easy to apply to general abelian Γ (and does not apply to non-abelian Γ at all). Consequently, a rather curious argument is employed in §3 to deal with the case of Γ abelian of odd order. This leads to a result which seems likely to be highly inaccurate, but is still adequate to give a positive answer to the above question.

The results here are by no means convincing evidence for the truth of the finiteness statement. However, a negative answer to the question would be of some interest. For, suppose we had infinitely many $x \in \mathbf{Z}\Gamma^\times$ such that $|\nu_a(x)| < M$. The numbers $\nu_a(x)$, $x \in \mathbf{Z}\Gamma^\times$, are non-zero rational numbers with bounded denominators, so there must be a positive number $M' < M$ such that $|\nu_a(y)| = M'$ for infinitely many $y \in \mathbf{Z}\Gamma^\times$. There are only finitely many possibilities for the principal fractional ideal $(a \cdot y)\mathfrak{o}_K$ of $\mathfrak{o}_K$ of absolute norm M', whence we can produce infinitely many units of $\mathfrak{o}_K$ of the form $(a \cdot y)/(a \cdot y')$.

1. Diophantine Approximation

Let V be a finite-dimensional real vector space. A *Euclidean norm* on V is a function $v \mapsto \|v\|$, $v \in V$, taking non-negative real values, such that

$\|v\| = 0$ *if and only if* $v = 0$;
$\|\lambda v\| = |\lambda|\|v\|$, *for all* $\lambda \in \mathbf{R}$, $v \in V$;
$\|v_1 + v_2\| \leqslant \|v_1\| + \|v_2\|$, *for all* $v_1, v_2 \in V$.

Any two such norms are bounded in terms of each other, and it will not matter which we choose. We tend to change frequently, using whichever seems most convenient at the time.

Let $l_1, l_2, \ldots, l_t$ be homogeneous linear forms on $\mathbf{R}^n$ with real algebraic coefficients. (One can also work with complex coefficients, but the gain in generality is minimal, for our purposes.) We say that a real subspace V of $\mathbf{R}^n$ is *defined over* $\mathbf{Q}$ if it has a basis over $\mathbf{R}$ consisting of elements of $\mathbf{Q}^n$. This is equivalent to V being spanned, over $\mathbf{R}$, by $V \cap \mathbf{Q}^n$.

THEOREM A (W. M. Schmidt, [3]). *Let $\eta > 0$. The following are equivalent:*

(a) *There is a constant $c > 0$ such that the inequality*

$$\left| \prod_{i=1}^{t} l_i(x) \right| \leqslant c\|x\|^{t-\eta}$$

has infinitely many solutions $x \in \mathbf{Z}^n$.

(b) *There is a set of indices $l \leqslant i_1 < i_2 < \cdots < i_m < t$, and a d-dimensional subspace V of $\mathbf{R}^n$, defined over $\mathbf{Q}$, such that the restrictions $l_{i_j} \mid V$, $1 \leqslant j \leqslant m$, of the l_{i_j} to V form a system of rank r, where these constants satisfy $r < d$, and $r \leqslant dm/\eta$.*

The rank condition in (b) means simply this. If we define a linear map $F: \mathbf{R}^n \to \mathbf{R}^m$ by

$$F(v) = (l_{i_1}(v), \ldots, l_{i_m}(v)),$$

then $\dim_{\mathbf{R}}(F(V)) = r$.

We shall also need, in §3, the following result:

THEOREM B [3]. *Let E be an algebraic number field, and M a finitely generated additive subgroup of E. Then there exists a constant c such that the equation $N_{E/\mathbf{Q}}(m) = c$ has infinitely many solutions $m \in M$ if and only if the $\mathbf{Q}$-vector space $\mathbf{Q}M$ spanned by M contains a set of the form αk, where $\alpha \in E^{\times}$, and k is a subfield of E which is neither $\mathbf{Q}$ nor an imaginary quadratic field.*

Before treating our norm distribution problem, we derive one fairly general consequence of Theorem A. For this, we let $E/\mathbf{Q}$ be a totally real algebraic number field, with $E/\mathbf{Q}$ Galois, and $E \neq \mathbf{Q}$. Let K be another real (but not necessarily totally real) number field, which is *linearly disjoint from E over $\mathbf{Q}$*. It is convenient to view E and K as subfields of $\mathbf{R}$. We let $m: E \to K$ be an *injective* homomorphism of $\mathbf{Q}$-vector spaces. We write $\mathfrak{o}_E$ for the ring of algebraic integers in E, and we choose a $\mathbf{Z}$-basis $\alpha_1, \ldots, \alpha_n$ of $\mathfrak{o}_E$, where $n = [E:\mathbf{Q}]$. We use this basis to identify $\mathfrak{o}_E$ with $\mathbf{Z}^n$, E with $\mathbf{Q}^n$, and $E \otimes_{\mathbf{Q}} \mathbf{R}$ with $\mathbf{R}^n$, as vector spaces. Then m extends by linearity to a homogeneous linear form on $E \otimes_{\mathbf{Q}} \mathbf{R} = \mathbf{R}^n$ with real algebraic coefficients, lying in K. These coefficients are linearly independent over $\mathbf{Q}$, by hypothesis, and indeed over E, by the linear disjointness condition.

For $\alpha \in E$, we have

$$N_{E/\mathbf{Q}}(\alpha) = \prod_{i=1}^{n} \sigma_i(\alpha),$$

where σ_i ranges over the distinct embeddings of E in $\mathbf{R}$. Identifying $E \otimes_{\mathbf{Q}} \mathbf{R}$ with $\mathbf{R}^n$, via our basis $\alpha_1, \ldots, \alpha_n$, we may again view the σ_i as linear forms on $\mathbf{R}^n$ with real algebraic coefficients, this time lying in E. Now we have:

PROPOSITION 1. *Use the above notation. Let $\|\cdot\|$ be a Euclidean norm on $E \otimes_{\mathbf{Q}} \mathbf{R} = \mathbf{R}^n$. Then, given $\varepsilon > 0$, there exists $c > 0$ such that*

$$|m(\alpha)N_{E/\mathbf{Q}}(\alpha)| \geqslant c\|\alpha\|^{1-\varepsilon} \quad \textit{for all } \alpha \in \mathfrak{o}_E.$$

In particular,

$$|m(\alpha)| \geqslant c\|\alpha\|^{1-\varepsilon} \quad \textit{for all } \alpha \in \mathfrak{o}_E^{\times}.$$

Proof. The second statement follows immediately from the first. To prove the first, we start by applying Theorem A to the linear forms σ_i, $1 \leqslant i \leqslant n$, on $\mathbf{R}^n = E \otimes_{\mathbf{Q}} \mathbf{R}$. We have

$$\left|\prod_{i=1}^{n} \sigma_i(\alpha)\right| \geqslant 1 = \|\alpha\|^0 \quad \text{for all } \alpha \in \mathfrak{o}_E, \alpha \neq 0.$$

The unit group $\mathfrak{o}_E^{\times}$ is infinite, so equality holds here for infinitely many $\alpha \in \mathfrak{o}_E$. Thus the maximum value of η for which Theorem A(a) holds, for this family of forms, is $n = [E:\mathbf{Q}]$.

Now we consider the family $S = \{m\} \cup \{\sigma_i : 1 \leqslant i \leqslant n\}$, and find the maximum value η_0 of η for which the statements of Theorem A hold. If T is a non-empty subset of S, with t elements, and V is a d-dimensional subspace of $\mathbf{R}^n$, defined over $\mathbf{Q}$, such that T has rank r on V, with $r < d$, we let

$$\eta(T, V) = dt/r.$$

Then

$$\eta_0 = \max_{T,V} \eta(T, V).$$

We know that if $m \notin T$, we have $\eta(T, V) \leqslant n$. We next take $T = \{m\}$. Then m has no non-trivial zeros on E (i.e. $\mathbf{Q}^n \subset \mathbf{R}^n$), so we must have $r = t = 1$, whence

$$\eta(T, V) = d \leqslant n.$$

Now we take the general case $T = \{m, \sigma_1, \ldots, \sigma_s\}$, $s \geqslant 1$, after renumbering the σ_i as convenient. Let r' be the rank of the system $\{\sigma_1 \mid V, \ldots, \sigma_s \mid V\}$.

Then $r' = r$ or $r - 1$. In the second case,

$$\eta(T, V) = d(s + 1)/(r' + 1) \leqslant ds/r',$$

since $r' \leqslant s$, and we have $\eta(T, V) \leqslant ds/r' \leqslant n$, by our earlier remark. In the first case, when $r' = r$, we must have a linear dependence relation

$$m \mid V = \sum_{i=1}^{e} \lambda_i \sigma_i \mid V,$$

with $\lambda_i \in \mathbf{R}$, after renumbering $\sigma_1, \ldots, \sigma_s$ again. We can assume that the $\sigma_i \mid V$, $1 \leqslant i \leqslant e$, are linearly independent. This forces $e \leqslant r < d$. The $\mathbf{Q}$-vector space $W_0 = V \cap E$ (i.e., $V \cap \mathbf{Q}^n$) has $\mathbf{Q}$-dimension d, by hypothesis on V. Therefore the E-vector space $W = E \cdot W_0$ (where we now regard E as a subfield of $\mathbf{R}$, acting on $\mathbf{R}^n = E \otimes_{\mathbf{Q}} \mathbf{R}$ via the *second* tensor factor), has E-dimension d also. We have seen that $e < d$, so the σ_i, $1 \leqslant i \leqslant e$, have a non-trivial common zero β on W. Then $\beta = \Sigma_{i=1}^{n} \mu_i \alpha_i$, where the $\mu_i \in E$, and we have $m(\beta) = 0 = \Sigma_{i=1}^{n} \mu_i m(\alpha_i)$. We know that the coefficients $m(\alpha_i)$ of m are linearly dependent over E, and we have a contradiction. Therefore only the case $r' = r - 1$ can arise, and we have $\eta(T, V) \leqslant n$. We can achieve $\eta(T, V) = n$ by taking $V = \mathbf{R}^n$, $|T| = 1$, and it follows that $\eta_0 = n$. The Proposition now follows from Theorem A.

2. Cyclic Extensions of Prime Degree

We return to the notation of the introduction, with $\Gamma = \mathrm{Gal}(K/\mathbf{Q})$, and we fix an element $a \in K$ such that $a \cdot \mathbf{Q}\Gamma = K$. We set

$$\nu_a(x) = N_{K/\mathbf{Q}}(a \cdot x), \quad x \in \mathbf{Q}\Gamma.$$

For each $\gamma \in \Gamma$, we let

$$l_\gamma(x) = \sum_{\delta \in \Gamma} a^{\delta\gamma} x_\delta \quad \text{where} \quad x = \sum_{\delta \in \Gamma} x_\delta \delta \in \mathbf{Q}\Gamma,$$

so that

$$\nu_a(x) = \prod_{\gamma \in \Gamma} l_\gamma(x).$$

The functions l_γ, $\gamma \in \Gamma$, are linear forms on $\mathbf{Q}\Gamma$, and we may use the same formula to view them as linear forms on $\mathbf{R}\Gamma$. We also let $\|\cdot\|$ denote a Euclidean norm on $\mathbf{R}\Gamma$, as at the beginning of §1.

THEOREM 1. *Suppose that Γ is cyclic of prime order p. Then, given $\varepsilon > 0$, there exists $c > 0$ such that $|\nu_a(x)| \geqslant c\|x\|^{p-\varepsilon}$, for all $x \in \mathbf{Z}\Gamma^{\times}$.*

Remark. With no hypothesis on Γ, elementary considerations show that there is a constant $c' > 0$ such that $|\nu_a(x)| \leqslant c'\|x\|^{|\Gamma|}$, for all $x \in \mathbf{Q}\Gamma$. For, $\nu_a(x)$ is a product of $|\Gamma|$ linear forms $l(x)$, each of which satisfies an inequality

$|l(x)| \leqslant c(l)\|x\|$, for all $x \in \mathbf{Q}\Gamma$, and some constant $c(l) > 0$. Thus Theorem 1 is best possible, of its type.

Proof. It is enough to prove the assertion when x is only allowed to range over a subgroup G of $\mathbf{Z}\Gamma^\times$ of finite index. To prove this, we choose a (finite) set of coset representatives x_i of $\mathbf{Z}\Gamma^\times/G$, and the theorem for G then gives constants $c_i > 0$ such that

$$|\nu_{ax_i}(x)| = |\nu_a(x_i x)| \geqslant c_i\|x\|^{p-\varepsilon},$$

for all $x \in G$. Multiplication by x_i is a linear automorphism of $\mathbf{R}\Gamma$, so there exist constants $b_i > 0$ such that $\|x_i x\| \leqslant b_i\|x\|$, for all $x \in \mathbf{R}\Gamma$. Taking c to be the minimum of the $c_i b_i^{\varepsilon-p}$, the result follows for $\mathbf{Z}\Gamma^\times$.

A suitable subgroup G is defined as follows. Pick a generator ξ of Γ, and a primitive *p-th* root of unity ζ. Then $\xi \mapsto (1, \zeta)$ induces an isomorphism $\mathbf{Q}\Gamma \simeq \mathbf{Q} \times \mathbf{Q}(\zeta)$ of $\mathbf{Q}$-algebras. Write π_0, π_1, for the projections $\mathbf{Q}\Gamma \to \mathbf{Q}$, $\mathbf{Q}\Gamma \to \mathbf{Q}(\zeta)$, respectively. We let G be the group of all $x \in \mathbf{Z}\Gamma^\times$ such that $\pi_0(x) = 1$, and $\pi_1(x)$ is a totally real unit in $\mathbf{Q}(\zeta)$. The Dirichlet unit theorem (in $\mathbf{Q}(\zeta)$) confirms that the index $(\mathbf{Z}\Gamma^\times : G)$ is finite. The fields K and $\mathbf{Q}(\zeta)$ are linearly disjoint over $\mathbf{Q}$, so we may apply Prop. 1 with $m = l_\gamma$, E the maximal totally real subfield of $\mathbf{Q}(\zeta)$, to obtain

$$|l_\gamma(\pi_1(x))| \geqslant c\|\pi_1(x)\|^{1-\varepsilon},$$

for all $\gamma \in \Gamma$ and all $x \in G$. But $l_\gamma(x) = l_\gamma(\pi_0(x)) + l_\gamma(\pi_1(x)) = \mathrm{Tr}_{K/\mathbf{Q}}(a) + l_\gamma(\pi_1(x))$, for all $x \in G$. Moreover, $\|\pi_1(x)\|/\|x\| \to 1$ as $\|x\| \to \infty$, $x \in G$. It follows that there is a constant $c > 0$ such that

$$|l_\gamma(x)| \geqslant c\|x\|^{1-\varepsilon},$$

for all $\gamma \in \Gamma$ and all $x \in G$. Since $\nu_a = \Pi\, l_\gamma$, the result follows.

Remark. The theorem holds, with essentially the same proof, but with p replaced by $|\Gamma|$ in the statement, when Γ is cyclic of order 8 or 9, provided one assumes that K is linearly disjoint from the field of $|\Gamma|$-*th* roots of unity. It is also easy to extend this to the case of Γ cyclic of order $2p$ (p prime), 16, or 18, when K is totally imaginary. The same applies to the following corollaries.

COROLLARY 1. *Let $M > 0$. Then, under the hypotheses of Theorem* 1, *the inequality $|\nu_a(x)| \leqslant M$ has only finitely many solutions $x \in \mathbf{Z}\Gamma^\times$.*

Proof. The theorem implies that there is a constant $M' > 0$ such that $\|x\| < M'$ for all $x \in \mathbf{Z}\Gamma^\times$ satisfying $|\nu_a(x)| < M$. There are only finitely many $y \in \mathbf{Z}\Gamma$ such that $\|y\| < M'$, and this proves the corollary.

COROLLARY 2. *Under the hypotheses of Theorem* 1, *let $\{u_n\}$, $n \geqslant 1$, be a sequence of distinct units of the ring $\mathfrak{o}_K$ of algebraic integers in K. Then,*

with at most finitely many exceptions, the group index $i_n = (\mathfrak{o}_K : (u_n \cdot \mathbf{Z}\Gamma))$ *is finite. Further,* $i_n \to \infty$ *as* $n \to \infty$.

Proof. For $b \in K$, we have $b \cdot \mathbf{Q}\Gamma = K$ if and only if $b \notin \mathbf{Q} \cup \mathrm{Ker}(\mathrm{Tr}_{K/\mathbf{Q}})$. The vector spaces $\mathbf{Q}$, $\mathrm{Ker}(\mathrm{Tr}_{K/\mathbf{Q}})$ only contain finitely many units, by Theorem B. For $b \in \mathfrak{o}_K$, the index $(\mathfrak{o}_K : b \cdot \mathbf{Z}\Gamma)$ is finite if and only if $b \cdot \mathbf{Q}\Gamma = K$. This proves the first assertion. Let us assume that the second assertion is false: there exists a sequence $\{u_n\}$ such that the index i_n is bounded. Since the values of i_n are positive integers, we may pass to a subsequence and assume that i_n is constant. The lattice $\mathfrak{o}_K$ has only finitely many sublattices of given finite index, and it follows that there is a value n_0 such that $u_{n_0} \cdot \mathbf{Z}\Gamma = u_m \cdot \mathbf{Z}\Gamma$, for infinitely many m. Then the Galois orbit $u_{n_0} \cdot \mathbf{Z}\Gamma^\times$ contains infinitely many units, which contradicts Corollary 1.

3. Abelian Extensions of Odd Degree

We continue with the notations introduced at the beginning of §2.

THEOREM 2. *Suppose that* Γ *is abelian of odd order* n *and exponent* e. *Suppose also that* K *is linearly disjoint over* $\mathbf{Q}$ *from the field of* e*-th roots of unity. Then, given* $\varepsilon > 0$, *there exists* $c > 0$ *such that*

$$|\nu_a(x)| \geqslant c\|x\|^{2/(n-1)-\varepsilon} \quad \text{for all } x \in \mathbf{Z}\Gamma^\times.$$

Proof. This is an awkward combination of Theorems A and B. Observe to start with that our number field K is totally real.

The rings $\mathbf{R}\Gamma$, $\mathbf{Q}\Gamma$, $\mathbf{Z}\Gamma$, all admit an involution "bar" which fixes coefficients, and satisfies $\bar{\gamma} = \gamma^{-1}$, $\gamma \in \Gamma$. We let $(\mathbf{R}\Gamma)_0$, $(\mathbf{Q}\Gamma)_0$, $(\mathbf{Z}\Gamma)_0$ denote the subrings of elements fixed by this involution. The ring $\mathbf{Q}\Gamma$ is a product of full cyclotomic fields $\mathbf{Q}(\zeta)$, where ζ is a root of unity of order dividing e, and $(\mathbf{Q}\Gamma)_0$ is the product of the maximal totally real subfields of these cyclotomic factors. The Dirichlet unit theorem implies that $(\mathbf{Z}\Gamma)_0^\times$ is of finite index in $\mathbf{Z}\Gamma^\times$, and we need only prove our theorem for x ranging over $(\mathbf{Z}\Gamma)_0^\times$. We first need to establish:

MAIN LEMMA. *Let* k *be a subfield of* K, $k \neq \mathbf{Q}$. *Then the* $\mathbf{Q}$*-vector space* $a \cdot (\mathbf{Q}\Gamma)_0 \subset K$ *contains no non-zero* k*-vector space.*

Proof. We work by induction on the order n of Γ. If Γ has prime order, the assertion is trivial, for reasons of dimension. So let us assume that Γ has composite order, and that there exists $\alpha \in K^\times$ such that $\alpha k \subset a \cdot (\mathbf{Q}\Gamma)_0$. We may replace k by a smaller field if possible, and assume that $[k:\mathbf{Q}] = p$ is prime. Let Δ be the subgroup of Γ fixing k. Now we choose a maximal proper subfield L of K containing k. Let Σ be the subgroup of

Δ which fixes L. Define

$$W = \{w \in a \cdot (\mathbf{Q}\Gamma)_0 \colon wk \subseteq a \cdot (\mathbf{Q}\Gamma)_0\}.$$

This is the unique maximal k-subspace of $a \cdot (\mathbf{Q}\Gamma)_0$. The element $\Sigma_{\sigma\in\Sigma}\,\sigma$ of $\mathbf{Q}\Gamma$ certainly lies in $(\mathbf{Q}\Gamma)_0$, and it acts on K as the relative trace $\mathrm{Tr}_{K/L}$. Therefore the space $a \cdot (\mathbf{Q}\Gamma)_0$ is invariant under the trace operator $\mathrm{Tr}_{K/L}$. Now take $w \in W$, and set $u = \mathrm{Tr}_{K/L}(w) \in a \cdot (\mathbf{Q}\Gamma)_0$. We may view the group algebra $\mathbf{Q}[\Gamma/\Sigma]$ as an ideal of $\mathbf{Q}\Gamma$, and it is a direct factor of $\mathbf{Q}\Gamma$. We have

$$\mathrm{Tr}_{K/L}(a \cdot \mathbf{Q}\Gamma) = \mathrm{Tr}_{K/L}(a) \cdot \mathbf{Q}[\Gamma/\Sigma],$$

where we view L as a Γ/Σ-module in the natural way. Moreover

$$\mathrm{Tr}_{K/L}(a \cdot (\mathbf{Q}\Gamma)_0) = \mathrm{Tr}_{K/L}(a) \cdot (\mathbf{Q}[\Gamma/\Sigma])_0.$$

This gives us

$$uk \subseteq \mathrm{Tr}_{K/L}(a) \cdot (\mathbf{Q}[\Gamma/\Sigma])_0.$$

But $\mathrm{Tr}_{K/L}(a) \cdot \mathbf{Q}[\Gamma/\Sigma] = L$, and if $u \neq 0$, the inductive hypothesis gives a contradiction. Therefore we must have

$$W \subseteq \mathrm{Ker}(\mathrm{Tr}_{K/L}).$$

The set W is a k-vector space, by its definition, and it is also a $(\mathbf{Q}\Delta)_0$-module. For, $(\mathbf{Q}\Delta)_0$ has a $\mathbf{Q}$-basis consisting of elements $\delta + \delta^{-1}$, $\delta \in \Delta$. If $\delta \in \Delta$, $\beta \in k$, $w \in W$, we have

$$(w^{\delta} + w^{\delta^{-1}})\beta = (w\beta) \cdot (\delta + \delta^{-1}) \in a \cdot (\mathbf{Q}\Gamma)_0 \cdot (\delta + \delta^{-1}) \subseteq a \cdot (\mathbf{Q}\Gamma)_0.$$

This shows moreover that W is a module over the algebra $k \otimes_{\mathbf{Q}} (\mathbf{Q}\Delta)_0$. Let $K\langle\Gamma\rangle$ denote the *twisted* group algebra of Γ over K. This is the left K-vector space with basis $\{\gamma \colon \gamma \in \Gamma\}$, and multiplication

$$\gamma_1 \cdot \gamma_2 = \gamma_1\gamma_2, \quad \gamma_i \in \Gamma,$$

$$\alpha \cdot \gamma = \alpha^{\gamma^{-1}}\gamma, \quad \alpha \in K, \gamma \in \Gamma.$$

Both k and $(\mathbf{Q}\Delta)_0$ are subalgebras of $K\langle\Gamma\rangle$ in a natural way. Together, they generate a subalgebra $k \cdot (\mathbf{Q}\Delta)_0$ which is isomorphic to $k \otimes_{\mathbf{Q}} (\mathbf{Q}\Delta)_0$. The action of $k \otimes_{\mathbf{Q}} (\mathbf{Q}\Delta)_0$ on W coincides with the action of $k \cdot (\mathbf{Q}\Delta)_0$ induced by the natural action of $K\langle\Gamma\rangle$ on K. Assuming $W \neq \{0\}$, it has a simple $k \cdot (\mathbf{Q}\Delta)_0$-submodule U, say. The linear disjointness condition on K then implies that, as $(\mathbf{Q}\Delta)_0$-module, U is isomorphic to a direct sum of $p = [k:\mathbf{Q}]$ copies of a simple $(\mathbf{Q}\Delta)_0$-module.

Now consider the $\mathbf{Q}\Gamma$-module $\mathrm{Ker}(\mathrm{Tr}_{K/L})$. Under the isomorphism $K \simeq \mathbf{Q}\Gamma$, $\mathrm{Ker}(\mathrm{Tr}_{K/L})$ corresponds to the kernel of the canonical projection π_Σ: $\mathbf{Q}\Gamma \to \mathbf{Q}[\Gamma/\Sigma]$. This is a direct sum of pairwise non-isomorphic *simple* $\mathbf{Q}\Gamma$-modules. On the other hand, it is the $\mathbf{Q}\Gamma$-module induced from the kernel of the canonical projection of $\mathbf{Q}\Delta$ onto $\mathbf{Q}[\Delta/\Sigma]$. For any simple $\mathbf{Q}\Delta$-module M, we have $M \otimes_{\mathbf{Q}\Delta} \mathbf{Q}\Gamma \simeq M^p$ as $\mathbf{Q}\Delta$-module. Therefore, as $\mathbf{Q}\Delta$-module,

we have $\mathrm{Ker}(\mathrm{Tr}_{K/L}) = \oplus_{i=1}^{r} M_i^p$, where each M_i is a simple $\mathbf{Q}\Delta$-module, and $M_i \neq M_j$ if $i \neq j$. The terms M_i^p are the simple $\mathbf{Q}\Gamma$-components of $\mathrm{Ker}(\mathrm{Tr}_{K/L})$.

Now we restrict scalars to $(\mathbf{Q}\Gamma)_0$, $(\mathbf{Q}\Delta)_0$, and consider $\mathrm{Ker}(\mathrm{Tr}_{K/L}) \cap a \cdot (\mathbf{Q}\Gamma)_0$. It is easy to see that we have the same structure. As $(\mathbf{Q}\Delta)_0$-module,

$$\mathrm{Ker}(\mathrm{Tr}_{K/L}) \cap a \cdot (\mathbf{Q}\Gamma)_0 = \bigoplus_{i=1}^{r} N_i^p,$$

where each N_i is a simple $(\mathbf{Q}\Delta)_0$-module, and $N_i \neq N_j$ if $i \neq j$. The terms N_i^p are the simple $(\mathbf{Q}\Gamma)_0$-components of

$$\mathrm{Ker}(\mathrm{Tr}_{K/L}) \cap a \cdot (\mathbf{Q}\Gamma)_0.$$

Our simple $k \cdot (\mathbf{Q}\Delta)_0$-submodule U of W ($\subseteq \mathrm{Ker}(\mathrm{Tr}_{K/L}) \cap a \cdot (\mathbf{Q}\Gamma)_0$) is of the form N^p, for some simple $(\mathbf{Q}\Delta)_0$-module N. Therefore $N^p = N_i^p$, for some i, and U is therefore a simple $(\mathbf{Q}\Gamma)_0$-component. In particular, U is a $(\mathbf{Q}\Gamma)_0$-module, and therefore a module over the subalgebra B of $K\langle\Gamma\rangle$ generated by k. $(\mathbf{Q}\Delta)_0$ and $(\mathbf{Q}\Gamma)_0$. Let $\gamma \in \Gamma$, $\gamma \notin \Delta$, and let $b \in k$ be some element which is not fixed by γ^2. Then

$$b(\gamma + \gamma^{-1}) = \gamma b^{\gamma} + \gamma^{-1} b^{\gamma^{-1}} \in B,$$

and of course $(\gamma + \gamma^{-1})b^{\gamma^{-1}} \in B$ also. Therefore $\gamma(b^{\gamma} - b^{\gamma^{-1}}) \in B$, and hence $\gamma \in B$. The elements $\gamma \in \Gamma$, $\gamma \notin \Delta$, generate Γ, and therefore $\Gamma \subset B$, $\mathbf{Q}\Gamma \subseteq B$. Thus U is a $\mathbf{Q}\Gamma$-module contained in $\mathrm{Ker}(\mathrm{Tr}_{K/L}) \cap a \cdot (\mathbf{Q}\Gamma)_0$. Passing to $\mathbf{Q}\Gamma$, we have constructed a non-trivial $\mathbf{Q}\Gamma$-module contained in $\mathrm{Ker}(\pi_\Sigma) \cap (\mathbf{Q}\Gamma)_0$. However, since Γ has odd order, the only $\mathbf{Q}\Gamma$-module contained in $(\mathbf{Q}\Gamma)_0$ is the direct factor $\mathbf{Q}$, corresponding to the trivial representation of Γ. This is certainly not contained in $\mathrm{Ker}(\pi_\Sigma)$. This contradiction proves the lemma.

Now we can apply Theorem B. We see that any equation

$$|\nu_a(x)| = c,$$

for c constant, has only finitely many solutions $x \in (\mathbf{Z}\Gamma)_0$. In our standard notation, we have

$$\nu_a(x) = \prod_{\gamma \in \Gamma} l_\gamma(x),$$

and we regard the l_γ as linear forms on $\mathbf{R}\Gamma$. We apply Theorem A to the forms $l_\gamma \mid (\mathbf{R}\Gamma)_0$, $\gamma \in \Gamma$. Let η_0 be the maximum value of η for which the statements of Theorem A hold for this family. We have just seen that we must have $\eta_0 < n = |\Gamma|$. But η_0 is the supremum of a set of rational numbers whose denominators r satisfy

$$r < \dim(\mathbf{R}\Gamma)_0 = (n + 1)/2 \quad \text{or} \quad r \leqslant (n - 1)/2.$$

Therefore $\eta_0 \leqslant n - 2/(n - 1)$, and the result follows from Theorem A.

Remark. We have actually proved that $|\nu_a(x)| > c\|x\|^{2/(n-1)-\varepsilon}$, for all $x \in (\mathbf{Z}\Gamma)_0$, so the discrepancy between Theorems 1 and 2 is only to be expected. It seems reasonable to expect that Theorem 2 can be greatly improved.

COROLLARY. *Under the hypotheses of Theorem* 2, *the inequality* $|\nu_a(x)| < M$ *has only finitely many solutions* $x \in \mathbf{Z}\Gamma^\times$, *for any given* $M > 0$.

To give a better picture of how the values $\nu_a(x)$, $x \in \mathbf{Z}\Gamma^\times$, are spread out, we include the following simple result, which shows that they are very thinly distributed indeed.

PROPOSITION 2. *Let* Γ *be a finite abelian group, and* $\|\cdot\|$ *a Euclidean norm on* $\mathbf{R}\Gamma$. *Then, if* s *is a complex variable, the series* $\Sigma_{x\in\mathbf{Z}\Gamma^\times} \|x\|^{-s}$ *converges for* $\mathrm{Re}(s) > 0$.

Proof. Let M be the unique maximal order in $\mathbf{Q}\Gamma$. It is enough to show that $\Sigma_{x\in\mathsf{M}^\times} \|x\|^{-s}$ converges for real $s > 0$. We may identify $\mathbf{Q}\Gamma$ with a finite direct product of algebraic number fields, when M becomes identified with the product of the rings of integers of these fields. Therefore it is certainly enough to prove convergence when M is replaced by any finite direct product:

$$\mathsf{M} = \prod_{i=1}^{n} \mathsf{O}_i,$$

where O_i is the ring of integers in a number field K_i, $1 \leqslant i \leqslant n$. The choice of Euclidean norm $\|\cdot\|$ on $(\Pi_{i=1}^n K_i) \otimes_{\mathbf{Q}} \mathbf{R}$ is clearly immaterial, so we take the following one. We let

$$\|(x_1, \ldots, x_n)\| = \max_{1\leqslant i\leqslant n} \|x_i\|, \quad x_i \in K_i \otimes_{\mathbf{Q}} \mathbf{R},$$

where the norm on $K_i \otimes_{\mathbf{Q}} \mathbf{R}$ is defined by identifying $K_i \otimes_{\mathbf{Q}} \mathbf{R}$ with the product of the Archimedean completions of K_i, and taking $\|x\|$ to be the absolutely largest coordinate. Thus, if $x \in K_i$, we have $\|x\| = \max_\sigma |\sigma(x)|$, where σ ranges over the embeddings of K_i in $\mathbf{C}$, and $|\cdot|$ is the ordinary absolute value on $\mathbf{C}$.

Let us first consider the case $n = 1$. The product formula for valuations of $K = K_1$ gives

$$\prod_\sigma |\sigma(x)| = 1 \quad \text{for all } x \in \mathsf{O}^\times,$$

where σ ranges over all distinct embeddings of K in $\mathbf{C}$, and $\mathsf{O} = \mathsf{O}_1$. It follows that

$$\min_\sigma |\sigma(x)| \geqslant \|x\|^{1-d} \quad \text{for all } x \in \mathsf{O}^\times,$$

where $d = [K:\mathbf{Q}]$. Now we define the usual logarithmic mapping

$$\mathscr{L}(x) = (\log|\sigma(x)|^{n(\sigma)}), \quad x \in \mathsf{O}^\times,$$

where σ now ranges over a set of representatives of the complex conjugacy classes of embeddings of K in $\mathbf{C}$, and $n(\sigma) = 1$ if $\sigma(K) \subset \mathbf{R}$, $n(\sigma) = 2$ otherwise. Then $\mathscr{L}$ identifies $\mathsf{O}^\times$ (mod roots of unity) with a lattice spanning a hyperplane in $\mathbf{R}^r$, r being the number of classes of embeddings σ. Taking the usual "absolutely largest coordinate" norm on $\mathbf{R}^r$, we have

$$\begin{aligned}\|\mathscr{L}(x)\| &= \max_\sigma \{n(\sigma)|\log|\sigma(x)\|\} \\ &= \max\{\max\{n(\sigma)\log|\sigma(x)|\}, -\min\{n(\sigma)\log|\sigma(x)|\}\}.\end{aligned}$$

However, $\max\{\log|\sigma(x)|\} \leqslant \log\|x\|$, and $\min\{\log|\sigma(x)|\} \geqslant (1 - d)\log\|x\|$, $x \in \mathsf{O}^\times$, so

$$\|\mathscr{L}(x)\| \leqslant 2(d - 1)\log\|x\|, \quad x \in \mathsf{O}^\times.$$

Therefore

$$\|x\| \geqslant \exp(\|\mathscr{L}(x)\|/2(d - 1)), \quad x \in \mathsf{O}^\times.$$

With our choice of norm on $K \otimes_{\mathbf{Q}} \mathbf{R}$, we have $\|x\| = \|\zeta x\|$, for any root of unity ζ. Thus we need only show that $\Sigma_{y \in \mathscr{L}(\mathsf{O}^\times)} \exp(-s\|y\|/2(d - 1))$ converges for real $s > 0$. But $\mathscr{L}(\mathsf{O}^\times) \simeq \mathbf{Z}^q$, for some q, and this reduces us to showing that $\Sigma_{z \in \mathbf{Z}^q} \exp(-s\|z\|)$ converges for $s > 0$, when $\|\cdot\|$ is any Euclidean norm on $\mathbf{R}^q$. The general case $n \geqslant 1$ reduces to exactly the same problem, by the same methods. Now we take $\|\cdot\|$ to be the absolutely largest coordinate norm on $\mathbf{R}^q$. The last series can be rearranged to $\Sigma_{N=0}^\infty \tau(N)\, e^{-sN}$, where $\tau(N)$ is the number of points $z \in \mathbf{Z}^q$ with $\|z\| = N$. Thus $\tau(N) = (2N + 1)^q - (2N - 1)^q$, which is $O(N^{q-1})$ as $N \to \infty$. Convergence is now immediate.

Remark. The convergence of the series $\Sigma_{x \in \mathbf{Z}\Gamma^\times} \|x\|^{-s}$ is not so spectacular when the finite group Γ is non-abelian. For quite trivial reasons, the abscissa of convergence is always less than or equal to $|\Gamma| - 1$. The first non-abelian example (with $|\Gamma| = 6$) has abscissa of convergence equal to 2.

4. The Universal Upper Bound

We continue to use the notations of the beginning of §2. We have already remarked that there is a constant c such that $|\nu_a(x)| \leqslant c\|x\|^{|\Gamma|}$, for all $x \in \mathbf{Q}\Gamma$. We conclude with a simple general result which shows that this bound cannot be reduced by restricting x to $\mathbf{Z}\Gamma^\times$.

THEOREM 3. *Let Γ be a finite group, abelian or non-abelian, such that $\mathbf{Z}\Gamma^\times$ is infinite. Then there is a constant $c > 0$ such that $|\nu_a(x)| \geqslant c\|x\|^{|\Gamma|}$, for infinitely many $x \in \mathbf{Z}\Gamma^\times$.*

Remark. The finite groups Γ for which $\mathbf{Z}\Gamma^{\times}$ is finite are listed in [2, §6].

Proof. Let $\mathbf{Q}\Gamma = \prod_{i=1}^{r} A_i$, where each A_i is a simple $\mathbf{Q}$-algebra. Write π_i for the canonical projection $\mathbf{Q}\Gamma \to A_i$. Let us assume first that one of the A_i, say A_1, is isomorphic to an algebra $M_k(D)$ of $k \times k$ matrices over some division algebra D, with $k \geqslant 2$. Then there exists an integer $n_0 \geqslant 1$ such that, for all $m \in \mathbf{Z}$, the element $x(m) \in \mathbf{Q}\Gamma$ determined by

$$\pi_1(x(m)) = \begin{bmatrix} 1 & mn_0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix},$$

$$\pi_i(x(m)) = 1 \quad \text{for } 2 \leqslant i \leqslant r,$$

actually lies in $\mathbf{Z}\Gamma$. Then $x(m) \in \mathbf{Z}\Gamma^{\times}$. For fixed $\gamma \in \Gamma$, the map $x \mapsto l_\gamma(x)$, $x \in \mathbf{Q}\Gamma$, is injective, by the choice of a. Therefore $l_\gamma(x(m))$ is a non-constant linear polynomial in m, and $l_\gamma(x(m))/m$ tends to a finite non-zero limit as $m \to \infty$. On the other hand, there are constants $c_1, c_2 > 0$ such that

$$c_1\|x(m)\| \leqslant m \leqslant c_2\|x(m)\|,$$

for all m. It follows that there is a constant $c' > 0$ such that $|l_\gamma(x(m))| \geqslant c'\|x(m)\|$, for all large m. Taking the product over $\gamma \in \Gamma$, we get the result in this case.

We are left with the case in which all the simple components A_i of $\mathbf{Q}\Gamma$ are division rings. Then, as in [2, p. 97], there is a simple component A_1, say, of $\mathbf{Q}\Gamma$ whose centre contains a totally real number field E with $[E:\mathbf{Q}] \geqslant 2$. Let $\sigma_1, \dots, \sigma_s$ be the distinct embeddings of E in $\mathbf{R}$. Then there is a unit u in E such that

$$|\sigma_1(u)| > 1, \ |\sigma_j(u)| < 1, \quad 2 \leqslant j \leqslant s.$$

This property is shared by any positive power of u, so we may choose u so that the element x of $\mathbf{Q}\Gamma$ defined by $\pi_1(x) = u$, $\pi_i(x) = 1$, $2 \leqslant i \leqslant r$, lies in $\mathbf{Z}\Gamma$. Then $x \in \mathbf{Z}\Gamma^{\times}$. Then $l_\gamma(x^m)$ is a linear polynomial in the conjugates $\sigma_j(u^m)$ of u^m, and so

$$|l_\gamma(x^m)| \geqslant c|\sigma_1(u)^m|,$$

with $c > 0$, for large m. Then, exactly as before, we get $|\nu_a(x^m)| \geqslant c\|x^m\|^{|\Gamma|}$, and the result.

References

1. C. J. Bushnell, *Norms of normal integral generators*, J. London Math. Soc. (2), vol. 15 (1977), pp. 191–209.
2. ———, *Norm distribution in Galois orbits*, J. reine angew. Math., vol. 310 (1979), pp. 81–99.

3. W. M. SCHMIDT, *Linearformen mit algebraischen Koeffizienten II,* Math. Ann., vol. 191 (1970), pp. 1–20.
4. M. J. TAYLOR, *On Fröhlich's conjecture for rings of integers in tame extensions,* Invent. Math., vol. 63 (1981), pp. 41–79.

UNIVERSITY OF LONDON KING'S COLLEGE
LONDON