Small zeros of hermitian forms over quaternion algebras

Lenny Fukshansky
Claremont McKenna College & IHES
(joint work with W. K. Chan)

Institut de Mathématiques de Jussieu
October 21, 2010

Cassels' Theorem

Let

$$F(\boldsymbol{X},\boldsymbol{Y}) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_{ij} X_i Y_j$$

be a symmetric bilinear form in $N \geq 2$ variables with coefficients in $\mathbb{Z}$. Write

$$F(\boldsymbol{X}) = F(\boldsymbol{X},\boldsymbol{X})$$

for the associated quadratic form.

In **1955, J. W. S. Cassels** proved that if F is isotropic, then there exists $\boldsymbol{x} \in \mathbb{Z} \setminus \{\boldsymbol{0}\}$ such that $F(\boldsymbol{x}) = 0$, and

$$(1) \qquad \max_{1 \leq i \leq N} |x_i| \leq \left(3 \sum_{i=1}^{N} \sum_{j=1}^{N} |f_{ij}| \right)^{\frac{N-1}{2}}.$$

The expressions on the left and right hand sides of (1) are examples of **heights** of a vector and of a quadratic form, respectively, and so (1) provides an explicit **search bound** for non-trivial zeros of F. The exponent $\frac{N-1}{2}$ in the upper bound is sharp, as shown by an example due to **M. Knesser**.

Over number fields

A generalization of Cassels' theorem over number fields has been obtained in **1975 by S. Raghavan**, who proved that if a quadratic form F with coefficients in a number field K is isotropic over K, then it has a non-trivial zero $\boldsymbol{x}$ over K of *small height*, where the bound is in terms of height of F for appropriately defined heights of $\boldsymbol{x}$ and F which generalize heights in (1). The exponent on height of F in the upper bound is again $\frac{N-1}{2}$.

Raghavan also produced an analogous result for zeros of **hermitian forms** over CM fields, where the exponent in the upper bound is $\frac{2N-1}{2}$. Although it is not clear whether this is sharp, it seems to be a correct analogue of the result for quadratic forms.

A more general result for quadratic forms was produced by J. D. Vaaler. To state it we need to develop some notation.

Notation and heights: number fields

Let K be a number field of degree d over $\mathbb{Q}$ with the set of places $M(K)$, and let O_K be its ring of integers.

For each $v \in M(K)$, let $d_v = [K_v : \mathbb{Q}_v]$ and let $| \ |_v$ be the unique absolute value on K_v that extends either the usual absolute value on $\mathbb{R}$ or $\mathbb{C}$ if $v|\infty$, or the usual p-adic absolute value on $\mathbb{Q}_p$ if $v|p$, where p is a rational prime.

Then the **product formula** reads:

$$\prod_{v \in M(K)} |a|_v^{d_v} = 1,$$

for each $0 \neq a \in K$.

Let $N \geq 2$, and define the following **infinite** and **finite heights** for each $\boldsymbol{x} \in K^N$:

$$H_{\mathsf{inf}}(\boldsymbol{x}) = \prod_{v|\infty} \max_{1 \leq i \leq N} |x_i|_v,$$

$$\mathcal{H}_{\mathsf{inf}}(\boldsymbol{x}) = \prod_{v|\infty} \sqrt{\sum_{i=1}^{N} |x_i|_v^2},$$

$$H_{\mathsf{fin}}(\boldsymbol{x}) = [O_K : O_K x_1 + \cdots + O_K x_N]^{-1}.$$

The **global heights** on K^N are defined by:

$$H(\boldsymbol{x}) = \left(H_{\mathsf{inf}}(\boldsymbol{x}) H_{\mathsf{fin}}(\boldsymbol{x})\right)^{\frac{1}{d}},$$

$$\mathcal{H}(\boldsymbol{x}) = \left(\mathcal{H}_{\mathsf{inf}}(\boldsymbol{x}) H_{\mathsf{fin}}(\boldsymbol{x})\right)^{\frac{1}{d}},$$

and the **inhomogeneous height** is given by

$$h(\boldsymbol{x}) := H(1, \boldsymbol{x}) \geq H(\boldsymbol{x}) \geq 1.$$

Due to the normalizing exponent $1/d$, our global height functions are **absolute**, i.e. for points over $\overline{\mathbb{Q}}$ their values do not depend on the field of definition, i.e. if $\boldsymbol{x} \in \overline{\mathbb{Q}}^N$ then height can be evaluated over any number field containing the coordinates of $\boldsymbol{x}$.

For a polynomial F with coefficients in K, $H(F)$ is the height of its coefficient vector.

For an L-dimensional subspace

$$V = \left\{ \boldsymbol{x} \in K^N : C\boldsymbol{x} = 0 \right\} \subseteq K^N,$$

where C is an $(N-L) \times N$ matrix of rank $N-L < N$ over K, let $\mathrm{Gr}(C)$ be the vector of Grassmann coordinates of C (i.e. determinants of $(N-L) \times (N-L)$ minors).

Then define

$$H_{\mathsf{inf}}(V) = \mathcal{H}_{\mathsf{inf}}(\mathsf{Gr}(C)).$$

Also define

$$H_{\mathsf{fin}}(V) = [O_K^{N-L} : C(O_K^N)]^{-1},$$

where C is viewed as a linear map

$$C : O_K^N \to O_K^{N-L}.$$

Then define

$$H(V) = (\mathcal{H}_{\mathsf{inf}}(V) H_{\mathsf{fin}}(V))^{\frac{1}{d}}.$$

This definition is independent of the choice of the basis by the product formula.

Vaaler's Theorem

In **1989, J. D. Vaaler** proved the following result. Let $V \subseteq K^N$ be an L-dimensional subspace, and let F be a quadratic form in N variables with coefficients in K which is isotropic over V. Then there exists a basis $\boldsymbol{x}_1, \dots, \boldsymbol{x}_L \in O_K^N$ for V such that $F(\boldsymbol{x}_i) = 0$ for all $1 \leq i \leq L$ with

$$h(\boldsymbol{x}_1) \leq h(\boldsymbol{x}_2) \leq \cdots \leq h(\boldsymbol{x}_L),$$

and

$$h(\boldsymbol{x}_1) h(\boldsymbol{x}_i) \ll_{K,N,L} H(F)^{L-1} H(V)^2, \tag{2}$$

for $1 \leq i \leq L$, where the constant in the upper bound is explicitly determined. In particular,

$$h(\boldsymbol{x}_1) \ll_{K,N,L} H(F)^{\frac{L-1}{2}} H(V), \tag{3}$$

which is the analogue of Cassels' bound, and is sharp at least with respect to the exponent on $H(F)$.

Our main result is an analogue of Vaaler's theorem over a certain class of quaternion algebras.

Notation and heights: quaternion algebras

Here we explain the notation and height machinery used in the statement of our result.

Let K as above be a **totally real** number field of degree d, and let $\alpha, \beta \in O_K$ be **totally negative**. Let $D = \left(\frac{\alpha,\beta}{K}\right)$ be a positive definite quaternion algebra over K, generated by the elements i, j, k which satisfy the following relations:

$$i^2 = \alpha,\ j^2 = \beta,\ ij = -ji = k,\ k^2 = -\alpha\beta.$$

As a vector space, D has dimension four over K, and $1, i, j, k$ is a basis. We fix this basis, and will always write each element $x \in D$ as

$$x = x(0) + x(1)i + x(2)j + x(3)k,$$

where $x(0), x(1), x(2), x(3) \in K$ are respective components of x, and the standard involution on D is conjugation:

$$\overline{x} = x(0) - x(1)i - x(2)j - x(3)k.$$

Trace and **norm** on D are defined by

$$\mathrm{Tr}(x) = x + \overline{x} = 2x(0),$$

$$\mathrm{N}(x) = x\overline{x} = x(0)^2 - \alpha x(1)^2 - \beta x(2)^2 + \alpha\beta x(3)^2$$

The algebra D is **positive definite** meaning that the norm $\mathrm{N}(x)$ is given by a positive definite quadratic form. In fact, since the norm form $\mathrm{N}(x)$ is positive definite,

$$D_v := D \otimes_K K_v$$

is isomorphic to the real quaternion

$$\mathbb{H} = \mathbb{R} + \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$$

for each $v \in M(K)$ such that $v|\infty$; there are d such places, corresponding to the embeddings of K, call them $v_1, \dots, v_d$. Then each embedding $\sigma_n : K \to \mathbb{R}$, $1 \leq n \leq d$, induces an embedding $\sigma_n : D \to D_{v_n}$, given by

$$\sigma_n(x) = x(0)^{(n)} + x(1)^{(n)} i + x(2)^{(n)} j + x(3)^{(n)} k$$

Write $x^{(n)}$ for $\sigma_n(x)$, then the **local norm**

$$\mathsf{N}^{(n)}(x) = x^{(n)}\overline{x}^{(n)}$$

at each archimedean place is also a positive definite quadratic form over the respective real completion K_{v_n}, $1 \leq n \leq d$. We now have archimedean absolute values on D, corresponding to the infinite places $v_1, \dots, v_d$ of K: for each $x \in D$, define

$$|x|_{v_n} = \sqrt{\mathsf{N}^{(n)}(x)}.$$

We can now define heights over D, following the work of **C. Liebendorfer, 2004**. First, we define the **infinite height** on D^N by

$$H_{\mathsf{inf}}(\boldsymbol{x}) = \prod_{n=1}^{d} \max_{1 \leq l \leq N} |x_l|_{v_n}$$

for every $\boldsymbol{x} \in D^N$.

Let us next fix an order $\mathcal{O}$ in D; our definition of **finite height** will be with respect to $\mathcal{O}$. For each $\boldsymbol{x} \in \mathcal{O}^N$, let

$$H_{\mathsf{fin}}^{\mathcal{O}}(\boldsymbol{x}) = [\mathcal{O} : \mathcal{O}x_1 + \cdots + \mathcal{O}x_N]^{-1/4}.$$

This is well defined, since $\mathcal{O}x_1 + \cdots + \mathcal{O}x_N$ is a left submodule of $\mathcal{O}$. Now we can define the **global homogeneous height** on $\mathcal{O}^N$ by

$$H^{\mathcal{O}}(\boldsymbol{x}) = \left(H_{\mathsf{inf}}(\boldsymbol{x}) H_{\mathsf{fin}}^{\mathcal{O}}(\boldsymbol{x})\right)^{1/d},$$

and the **global inhomogeneous height** by

$$h(\boldsymbol{x}) := H_{\mathsf{inf}}(1, \boldsymbol{x}) \geq H^{\mathcal{O}}(\boldsymbol{x}),$$

since $\mathcal{O} + \mathcal{O}x_1 + \cdots + \mathcal{O}x_N = \mathcal{O}$. To extend this definition to D^N, notice that for each $\boldsymbol{x} \in D^N$ there exists $a \in O_K$ such that $a\boldsymbol{x} \in \mathcal{O}^N$, and define $H^{\mathcal{O}}(\boldsymbol{x})$ to be $H^{\mathcal{O}}(a\boldsymbol{x})$ for any such a. This is well defined by the product formula, and $H^{\mathcal{O}}(\boldsymbol{x}t) = H^{\mathcal{O}}(\boldsymbol{x})$ for all $t \in D^{\times}$.

We can now define height on the set of proper right D-subspaces of D^N.

D splits over $E = K(\sqrt{\alpha})$, meaning that there exists a K-algebra homomorphism

$$\rho : D \to \mathrm{Mat}_{22}(E),$$

given by

$$\rho(x) = \begin{pmatrix} x(0) + x(1)\sqrt{\alpha} & x(2) + x(3)\sqrt{\alpha} \\ \beta(x(2) - x(3)\sqrt{\alpha}) & x(0) - x(1)\sqrt{\alpha} \end{pmatrix},$$

so that $\rho(D)$ spans $\mathrm{Mat}_{22}(E)$ as an E-vector space. This map extends naturally to matrices over D.

Let $Z \subseteq D^N$ be an L-dimensional right vector subspace of D^N, $1 \leq L < N$. Then there exists an $(N-L) \times N$ matrix C over D with left row rank $N - L$ such that

$$Z = \{\boldsymbol{x} \in D^N : C\boldsymbol{x} = \boldsymbol{0}\}.$$

Define

$$H_{\mathsf{inf}}(C) = \left(\prod_{n=1}^{d} \sum_{C_0} |\det(\rho(C_0))|_{v_n}^2\right)^{1/2},$$

where the sum is over all $(N-L) \times (N-L)$ minors C_0 of C.

Also define

$$H_{\mathsf{fin}}^{\mathcal{O}}(C) = [\mathcal{O}^{N-L} : C(\mathcal{O}^N)]^{-1/4},$$

where $C : \mathcal{O}^N \to \mathcal{O}^{N-L}$ is viewed as a linear map.

Then we can define

$$H^{\mathcal{O}}(Z) = \left(H_{\mathsf{inf}}(C) H_{\mathsf{fin}}^{\mathcal{O}}(C)\right)^{1/d}.$$

This definition does not depend on the specific choice of such matrix C.

For a polynomial F over D, heights of F are heights of its coefficient vector.

W. K. Chan, L.F. (Acta Arith., 2010)

Let $D = \left(\frac{\alpha,\beta}{K}\right)$ be a positive definite quaternion algebra over a totally real number field K, where α, β are totally negative algebraic integers in K. Let $\mathcal{O}$ be an order in D. Let $N \geq 2$ be an integer, and let $Z \subseteq D^N$ be an L-dimensional right D-subspace, $1 \leq L \leq N$. Let $F(\boldsymbol{X}, \boldsymbol{Y}) \in D[\boldsymbol{X}, \boldsymbol{Y}]$ be a hermitian form in $2N$ variables, and assume that F is isotropic on Z. Then there exists a basis $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L$ for Z over D such that

$$F(\boldsymbol{y}_n) := F(\boldsymbol{y}_n, \boldsymbol{y}_n) = 0$$

for all $1 \leq n \leq L$ and

(4) $h(\boldsymbol{y}_1) \ll_{K,\mathcal{O},N,L,\alpha,\beta} H_{\mathrm{inf}}(F)^{\frac{4L-1}{2}} H^{\mathcal{O}}(Z)^4,$

and

(5)

$h(\boldsymbol{y}_1) h(\boldsymbol{y}_n) \ll_{K,\mathcal{O},N,L,\alpha,\beta} H_{\mathrm{inf}}(F)^{4L-1} H^{\mathcal{O}}(Z)^8,$

where the constants in the upper bounds of (4) and (5) are explicitly determined.

Idea of the proof

Define a K-vector space isomorphism

$$[\ ]: D \to K^4,$$

given by

$$[x] = (x(0), x(1), x(2), x(3)),$$

for each $x = x(0) + x(1)i + x(2)j + x(3)k \in D$, which extends naturally to $[\ ]: D^N \to K^{4N}$, given by $[\boldsymbol{x}] = ([x_1], \dots, [x_N])$ for each $\boldsymbol{x} = (x_1, \dots, x_N) \in D^N$.

Define the **trace form**

$$Q([\boldsymbol{X}]) := \mathrm{Tr}(F(\boldsymbol{X}, \boldsymbol{X})),$$

which is a quadratic form in $4N$ variables over K. Then $F(\boldsymbol{x}) = 0$ for some $\boldsymbol{x} \in D^N$ if and only if $Q([\boldsymbol{x}]) = 0$.

Then apply Vaaler's theorem to $Q([\boldsymbol{X}])$ on the $4L$-dimensional subspace $[Z] \subseteq K^{4N}$. For this method to produce our result, we develop collection of height comparison lemmas between heights over K and heights over D.

Height comparison lemmas

First we compare the height of a vector $\boldsymbol{x} \in \mathcal{O}^N$ over D with its image in K^{4N} under the map [ ].

Lemma 1. *For each $\boldsymbol{x} \in \mathcal{O}^N$,*

$$\begin{aligned} t(\alpha,\beta) H([\boldsymbol{x}]) &\leq H_{\mathsf{inf}}(\boldsymbol{x}) \\ &\leq h(\boldsymbol{x}) \leq 2 s(\alpha,\beta) h([\boldsymbol{x}]), \end{aligned}$$

where

$$s(\alpha,\beta) = \prod_{n=1}^{d} \max\{1, |\alpha|_{v_n}, |\beta|_{v_n}, |\alpha\beta|_{v_n}\}^{\frac{1}{2}},$$

and

$$t(\alpha,\beta) = \prod_{n=1}^{d} \min\{1, |\alpha|_{v_n}, |\beta|_{v_n}, |\alpha\beta|_{v_n}\}^{\frac{1}{2}}.$$

Next we have the comparison for heights of the hermitian form F over D and its trace form Q over K.

Lemma 2. *Let F be a hermitian form over D and let Q be its associated trace form over K, as above. Then*

$$\frac{t(\alpha,\beta)}{2s(\alpha,\beta)^2} H(Q) \leq H_{\mathrm{inf}}(F),$$

and

$$H^{\mathcal{O}}(F) \leq 2^{\frac{d+1}{d}} s(\alpha,\beta) \mathbb{N}(\alpha\beta)^{\frac{1}{d}} \mathfrak{N}(\mathcal{O}) H(Q),$$

where $\mathbb{N}$ stands for the norm on K, and

$$\begin{aligned} \mathfrak{N}(\mathcal{O}) \;=\; \min\Big\{ & |\mathbb{N}(\gamma)|^{\frac{1}{d}} : \\ & \gamma \in O_K \text{ is such that } \gamma i, \gamma j, \gamma k \in \mathcal{O} \Big\}. \end{aligned}$$

In the next lemma we compare the heights of a D-subspace of D^N with respect to two different orders, $\mathcal{O}_1$ and $\mathcal{O}_2$ in D.

Lemma 3. *Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two orders in D, and let $Z \subseteq D^N$ be an L-dimensional right vector D-subspace of D^N, $1 \leq L \leq N$. Then*

$$\mathcal{M}^{-(N-L)} H^{\mathcal{O}_1}(Z) \leq H^{\mathcal{O}_2}(Z) \leq \mathcal{M}^{N-L} H^{\mathcal{O}_1}(Z),$$

where $\mathcal{M} = \mathcal{M}(\mathcal{O}_1, \mathcal{O}_2)$ is defined as

$$\mathcal{M} = \max \left\{ \mathbb{N}\left(\Delta_{\mathcal{O}_1} \Delta_{\mathcal{O}_2}^{-1}\right)^{\frac{1}{2}}, \mathbb{N}\left(\Delta_{\mathcal{O}_2} \Delta_{\mathcal{O}_1}^{-1}\right)^{\frac{1}{2}} \right\}.$$

Here $\Delta_{\mathcal{O}}$ is the discriminant of the order $\mathcal{O}$, which is the ideal in O_K generated by all the elements of the form

$$\det\left(\mathrm{Tr}(\omega_h \omega_n)\right)_{0 \leq h,n \leq 3} \in O_K,$$

where $\omega_0, \dots, \omega_3$ are in $\mathcal{O}$.

Finally, Lemma 3 can be used, along with the following identity, to relate height of a D-subspace of D^N to its image under [ ].

Lemma 4. *Let Z be as in Lemma 3, then $V_Z := [Z] \subseteq K^{4N}$ is a $4L$-dimensional K-subspace of K^{4N}, and*

$$H(V_Z) = H^{O_D}(Z)^4,$$

where

$$O_D = O_K + O_K i + O_K j + O_K k.$$

Remarks

The height $H_{\mathsf{inf}}(F)$ in our inequalities cannot be replaced by the height $H^{\mathcal{O}}(F)$: there are specific examples where the bounds would no longer hold.

While it is not clear whether our bounds are optimal, it is worth a note that starting from our bounds and applying our height comparison lemmas, one retrieves a Cassels-type bound with correct exponents for the corresponding quadratic trace form over K in $4N$ variables.

Moreover, our exponent on $H_{\mathsf{inf}}(F)$ in (4) is completely analogous to the bound obtained by Raghavan for hermitian forms over CM fields.