On the theorem[1] of Hasse-Minkowski

R. SRIDHARAN[2]

Let K be a field (which, in what follows, will always be assumed to be of characteristic different from two). A *quadratic form* in n variables is a homogeneous polynomial f in the variables $X_1, X_2, \cdots, X_n$:

$$f(X_1, X_2, \cdots, X_n) = \sum_{1 \leq i \leq n} a_i {X_i}^2 + \sum_{1 \leq i < j \leq n} 2a_i a_j X_i X_j \qquad (*)$$

An *isotropy* of a quadratic form over K is a non-zero vector, $(x_1, x_2, \cdots, x_n) \in K^n$, such that $f(x_1, x_2, \cdots, x_n) = 0$. We say that a quadratic form f over K is isotropic, if it has an isotropy.

Example: The quadratic form $X_1^2 - X_2^2$ (called the *hyperbolic plane*) is obviously isotropic. But the form $X_1^2 + X_2^2$ over the field $\mathbb{Q}$ of rational numbers is not isotropic.

Recall that an *algebraic number field* is by definition, a field extension of $\mathbb{Q}$ of finite degree. The classical theorem whose proof we wish to discuss here is due to Hasse-Minkowski, which gives a necessary and sufficient condition for a quadratic form over such a field to be isotropic in terms of "local" conditions. To state and prove this theorem we need some preliminary notation and results.

Let K be any field. Recall that a valuation of K is a map $v : K \to \mathbb{R}^+$, $\mathbb{R}^+$ denoting the non-negative real numbers, such that for $x, y \in K$,

1. $v(x) = 0$ if and only if $x = 0$,
2. $v(x + y) \leq v(x) + v(y)$,
3. $v(xy) = v(x)v(y)$, for $x, y \in K$.

[1]For some interesting historical comments on Hasse's proof of this theorem see the article of Hasse entitled "Kurt Hensels entscheidender Anstoss zur Entdeckung des Lokal-Global-Prinzips" in Crelle's Journal, 209, $3-4$, 1962, where Hasse discusses the important role played by Hensel in his proof of this theorem. Minkowski's proof was never published.

[2]I thank S. A. Katre and Dinesh Thakur for inviting me to give a few lectures on this theorem in the summer school on Cyclotomic Fields held at Pune in June 1999. I want take this opportunity to thank S. D. Adhikari who was mainly responsible for my writing up these notes. I am grateful to Preeti Raman who among other things, made this Tex version possible. As ever, I am deeply indebted to Parimala for all her help.

The map $v : K \to \mathbb{R}^+$ defined by $v(0) = 0$ and $v(x) = 1$ for $x \neq 0$ is obviously a valuation of K called *trivial valuation.* From now on by a valuation we mean a non-trivial valuation. We have a notion of equivalence of valuations defined by: $v \sim v'$, if and only if there exists a positive real number c such that $v' = v^c$. It is easy to see that v and v' are equivalent if and only if for any $x \in K$, $v(x) < 1$ implies $v'(x) < 1$. We shall often confuse between a valuation and its equivalence class. For any valuation v of K, we denote by K_v, the completion of K, with respect to v: K_v is in fact the completion of K with respect to the metric defined for $x, y \in K$ by $d(x, y) = v(x - y)$. The topological space K_v has an obvious field structure and K sits in K_v as a dense subfield.

We start with the following lemma on valuations.

Lemma.1: *Let $v_1, v_2, \cdots, v_n$ be pairwise inequivalent non-trivial valuations of any field K. Then the image of K under the canonical injection $K \to \prod_{1\le i\le n} K_{v_i}$ is dense. In other words, given $(x_1, x_2, \cdots, x_n) \in \prod_{1\le i\le n} K_{v_i}$ and an $\epsilon > 0$ in $\mathbb{R}$, there exists $x \in K$, such that $v_i(x - x_i) < \epsilon$.*

Proof. We note first that it is enough to find for each r, $1 \le r \le n$, $\theta_r \in K$ such that $v_r(\theta_r) > 1$, $v_m(\theta_r) < 1$, for $m \neq r$, $1 \le m \le n$. For, then as $s \to +\infty$, we have $\frac{\theta_r^s}{1+\theta_r^s} = \frac{1}{1+\theta_r^{-s}} \to 1$ with respect to v_r and $\frac{\theta_r^s}{1+\theta_r^s} = \frac{1}{1+\theta_r^{-s}} \to 0$ with respect to v_m, for $m \neq r$.Then it is enough to take $\xi = \sum_{r=1}^{n} \frac{\theta_r^s}{1+\theta_r^s} x_r$, for a sufficiently large s.

We show the existence of $\theta = \theta_1$, with $v_1(\theta) > 1$ and $v_r(\theta) < 1$ for $2 \le r \le n$. To do this we use induction on n.

Let $n = 2$. Since v_1 and v_2 are inequivalent, there exist α, β such that $v_1(\alpha) < 1$ and $v_2(\alpha) \ge 1$ and $v_1(\beta) \ge 1$ and $v_2(\beta) < 1$. Then $\theta = \beta\alpha^{-1}$ will do.

Let $n \ge 3$. By induction, there is a $\phi \in K$, such that $v_1(\phi) > 1$ and $v_r(\phi) < 1$, for $2 \le r \le n - 1$. By the case $n = 2$, there is a $\psi \in K$, such that $v_1(\psi) > 1$ and $v_n(\psi) < 1$. Then put

$$\theta = \begin{cases} \phi, & \text{if } v_n(\phi) < 1 \\ \phi^s \psi, & \text{if } v_n(\phi) = 1 \\ \frac{\phi^s}{1+\phi^s}\psi, & \text{if } v_n(\phi) > 1 \end{cases}$$

where $s \in \mathbb{N}$ is sufficiently large. This completes the proof.

We recall that the valuations of the field $\mathbb{Q}$ are given up to equivalence by the usual absolute value (called the *archimedean valuation* of $\mathbb{Q}$), or by the normalised p-adic valuation, corresponding to any prime p, defined for $a \neq 0$ by $v_p(a) = p^{-w_p(a)}$ and $v_p(0) = 0$, $p^{w_p(a)}$ being the maximum power of p which divides a. Let K be an algebraic number field. It is well known that any valuation of $\mathbb{Q}$ extends to finitely many inequivalent valuations of K. We note that if v is any valuation of K which is an extension of a p-adic valuation of $\mathbb{Q}$, then v satisfies the stronger condition $2'$: $v(x+y) \leq max(v(x), v(y))$, for $x, y \in K$.

If f is a quadratic form over K, which is isotropic, then obviously for any v, f is isotropic over K_v. The theorem of Hasse-Minkowski is indeed the converse of this statement, namely,

Theorem.2: *Let f be a quadratic form over an algebraic number field K. If f is isotropic over K_v for all v, then f is isotropic over K.*

In order to prove the theorem, we begin with some general facts on quadratic forms. Let f be a quadratic form, given by $f = \sum_{1 \leq i \leq n} a_i X_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} X_i X_j$. If we denote by $\bar{X}$, the row vector $(X_1, X_2, ..., X_n)$ and by A_f, the symmetric matrix whose diagonal entries are a_i and the off-diagonal $(i,j)^{th}$ entries are a_{ij}, then we have $f(\bar{X}) = \bar{X} A_f \bar{X}^t$. Let f and g be quadratic forms, given by $f = \sum_{1 \leq i \leq n} a_i X_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} X_i X_j$ and $g = \sum_{1 \leq i \leq n} b_i X_i^2 + \sum_{1 \leq i < j \leq n} 2b_{ij} X_i X_j$. We say that f and g are *equivalent*, written $f \sim g$ if there exists a $u \in Gl_n(K)$, such that $g(\bar{X}) = f(\bar{X}u)$. In other words, $\bar{X} A_g \bar{X}^t = \bar{X} u A_f (\bar{X}u)^t = \bar{X} u A_f u^t \bar{X}^t$. This implies that $A_g = u A_f u^t$ and conversely, if there exists a $u \in Gl_n(K)$, by reversing the steps we see that $f \sim g$. If f and g are equivalent and $A_g = u A_f u^t$, then $det(A_g) = det(A_f).det(u)^2$, so that the class of $det(A_f)$ modulo K^{*2} depends only on the equivalence class of f, and is called the *discriminant* of the quadratic form f, denoted by $disc(f)$. We say that a quadratic form f is *diagonal* if $a_{ij} = 0$ for $i \neq j$ or what is the same, A_f is a diagonal matrix. If the diagonal entries are $a_1, a_2, \cdots, a_n$, we shall denote in what follows, such a form by $< a_1, a_2, \cdots, a_n >$.

Proposition.3: *Any quadratic form (over a field of characteristic different from 2), is equivalent to a diagonal form.*

Proof. Let f be a quadratic form. Since the characteristic of K is not 2 and f is not identically zero, there exists $u = (u_1, u_2, \cdots, u_n) \in K^n$ such that $a = f(u_1, u_2, \cdots, u_n) = u A_f u^t \neq 0$. Let $W \subset K^n$ be defined by $W = \{w \in K^n \mid u A_f w^t = 0\}$. Obviously W is a subspace of K^n and $W \cap K.u = 0$, since $u \notin W$. We have that $K^n = W \oplus K.u$, since any $z \in K^n$ can be written as $(z - \lambda u) + \lambda u$ and for $\lambda = a^{-1} u A z^t$, $z - \lambda u \in W$. We now choose a basis of K^n which consists of u and a basis of W. For this choice of a basis of K^n, A_f has the form

$$\begin{pmatrix} a & 0 \\ 0 & B_{(n-1)\times(n-1)} \end{pmatrix}.$$

The proof now follows by induction on n.

In view of the above proposition, from now on, we assume that f is a diagonal form. The number of non-zero a_i (which is simply the rank of the matrix A_f), is called the *rank of f*.

We now prove the following fact, which will be used in the proof of the next proposition.

Lemma.4: *Any quadratic form of rank greater than or equal to 3, over a finite field of characteristic different from 2, is isotropic.*
Proof. [3] Obviously it is enough to show that the equation $aX^2+bY^2=1$ has a solution over any finite field $\mathbb{F}_q$ of q elements. The number of elements of the form $S=\{a\lambda^2 \mid \lambda\in\mathbb{F}_q\}$ has cardinality $\frac{q+1}{2}$, which is also the cardinality of the set $S'=\{1-b\mu^2 \mid \mu\in\mathbb{F}_q\}$. Since the number of elements of $\mathbb{F}_q$ is q, S and S' must intersect, which proves the lemma.

Let K be an algebraic number field, v a valuation of K and $F=K_v$ denote the completion of K at v. If v lies over the unique archimedean valuation of $\mathbb{Q}$, then K_v is isomorphic to either the real number field $\mathbb{R}$, or the field $\mathbb{C}$ of complex numbers. Assume now that v is non-archimedean. The set $\mathcal{O}_F=\{x\in F \mid v(x)\le 1\}$ is easily checked to be a subring of K and has a unique non-zero prime ideal, i.e., $\{x\in\mathcal{O}_F \mid v(x)<1\}$, which is principal. Any generator $\pi=\pi_v$ of this ideal is called a *uniformising parameter for v*. The field $\bar{F}=\mathcal{O}_F/(\pi)$ (called the *residue field at v*) is a finite extension of the prime field $\mathbb{Z}/p\mathbb{Z}$, (where v is an extension of the p-adic valuation of $\mathbb{Q}$) whose degree is denoted by f_v *(called the residue class field degree at v)*, so that $\bar{F}$ is a finite field with $q=p^{f_v}$ elements. By a *unit* of F, we mean an invertible element of $\mathcal{O}_F$, i.e., an element not in $(\pi\mathcal{O}_F)$. We note that an element $u\in F$ is a unit if and only if $v(u)=1$. We record the next proposition which is needed in the proof of theorem 2.

Proposition.5: *Let v be a non-dyadic, non-archimedean valuation of K, i.e., v is not an extension of the 2-adic valuation of $\mathbb{Q}$. Let $f=<u_1,u_2,u_3>$ be a rank 3 quadratic form over $F=K_v$, where u_i for $1\le i\le 3$, are units of F. Then f is isotropic over F.*

[3] I thank Dinesh Thakur for bringing to my attention the slick proof given here, which is different from the one I had given in the lecture.

Proof. In fact proving the proposition is equivalent to showing that there exist $\lambda_i \in \mathcal{O}_F$, for $1 \leq i \leq 3$, not all zero such that $\sum_{1\leq i\leq 3} u_i\lambda_i^2 = 0$. Since $\mathcal{O}_F$ is a subring of F, which is closed in F, it is complete for the topology of F. In fact $\mathcal{O}_F$ is a topological ring for which a fundamental system of neighbourhoods of 0 are (π^n) for $n \geq 0$.

Let $^-$ denote reduction modulo (π). Since $\bar{F}$ is a finite field, $\bar{f}$ is isotropic, by Lemma 4. Therefore there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathcal{O}_F$, such that $\sum_{1\leq i\leq 3} \bar{u_i}\bar{\lambda_i^2} = 0$. We can assume that one of the λ_i is a unit, let it be λ_1. We assume by induction that there exists an integer $n \geq 1$ and $\lambda_1, \lambda_2, \lambda_3$ in $\mathcal{O}_F$, with λ_1 a unit such that $\sum_{1\leq i\leq 3} u_i\lambda_i^2 \in (\pi^n)$, so that its image in $\mathcal{O}_F/(\pi^n)$ is zero. Certainly this is true for $n = 1$, as we have just now remarked. We wish to solve for an $h \in (\pi^n)$, such that $u_1(\lambda_1 + h)^2 + u_2\lambda_2^2 + u_3\lambda_3^2 \in (\pi^{n+1})$. Since $h \in (\pi^n)$, $h^2 \in (\pi^{2n}) \subset (\pi^{n+1})$, the choice $h = \frac{-(u_1\lambda_1^2+u_2\lambda_2^2+u_3\lambda_3^2)}{2u_1\lambda_1}$ would do. Since $\mathcal{O}_F$ is complete, an iteration of this procedure leads to a Cauchy sequence, which, since $\mathcal{O}_F$ is complete, converges and yields a solution of the equation $\sum_{1\leq i\leq 3} u_iX_i^2 = 0$ in $\mathcal{O}_F$. This proves the proposition.

The proof of the Hasse-Minkowski theorem is achieved by settling it for low ranks and then appealing to induction. To prove the theorem, we assume that $f =< a_1, a_2, \cdots, a_r >$, $r \geq 2$, and $a_i \neq 0$, for $1 \leq i \leq r$.

Let $r = 2$. Since $< a_1, a_2 > \ = a_1 < 1, a_1^{-1}a_2 >$, we may assume that $f =< 1, -\lambda >$, $\lambda \in K^*$, which is isotropic if and only if λ is a square in K. Theorem 2 in this case follows from the following theorem, which is a consequence of the first inequality of Class field theory.

Theorem.6: *Let K be a number field, $\lambda \in K$ and $L = K(\sqrt{\lambda})$. If for every valuation v of K, $K_v(\sqrt{\lambda}) = K_v$, then $K(\sqrt{\lambda}) = K$.*

Let $r = 3$. As above, we may assume that $f \ = \ < -1, \lambda, \mu >$. This form is isotropic if and only if there exist $\alpha_1, \alpha_2, \alpha_3 \in K$ such that $\alpha_1^2 = \lambda\alpha_2^2 + \mu\alpha_3^2$. Obviously α_2 and α_3 cannot both be zero. We may assume without loss in generality that α_2 is not zero, so that $\lambda = \frac{\alpha_1}{\alpha_2}^2 - \mu\frac{\alpha_3}{\alpha_2}^2$, i.e., λ is a norm in the extension $K(\sqrt{\mu})$ over K. In this case, the theorem of Hasse-Minkowski follows from the following more general,

Theorem.7: *(Hasse norm theorem) Let K be an algebraic number field and L over K a cyclic extension of K. Then an element $\lambda \in K$ is a norm from K if and only if $\lambda \in K_v$ is a norm from L_w over K_v, for all v, where w is some valuation of L, extending v.*

The theorem applied to the case where L over K is a quadratic extension yields the Hasse-Minkowski theorem, for $r = 3$.

Let $r = 4$. We may assume, as before by scaling that $f = < 1, -a, -b, c >$. We first consider the case where the discriminant of f is 1, i.e., $abc \in K^{*2}$, so that f can be replaced by the equivalent quadratic form $< 1, -a, -b, ab >$. The condition that f is isotropic is equivalent to saying that there exist $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in K$ such that ${\lambda_1}^2 - a{\lambda_2}^2 - b{\lambda_3}^2 + ab{\lambda_4}^2 = 0$. Let v be any valuation of K. If ${\lambda_3}^2 - a{\lambda_4}^2 = 0$, for some $\lambda_3, \lambda_4 \in K_v^*$, then $K_v(\sqrt{a}) = K_v$, so that any element of K_v, in particular b is a norm from $K_v(\sqrt{a}) = K_v$. If on the other hand there is a solution in K_v, with ${\lambda_3}^2 - a{\lambda_4}^2 \neq 0$, then $b = \frac{{\lambda_1}^2 - a{\lambda_2}^2}{{\lambda_3}^2 - a{\lambda_4}^2}$ is a norm in $K_v(\sqrt{a})$ over K_v. The Hasse norm theorem quoted above shows then that $b \in K$ is a norm from $K(\sqrt{a})$. This implies that $b = {\lambda_1}^2 - a{\lambda_2}^2$, which shows that f is isotropic over K.

Let now $f =< 1, -a, -b, c >$ and $disc(f) = abc \notin K^{*2}$. To prove the theorem in this case, we need the following lemma.

Lemma.8: *Let $q =< a_1, a_2, \cdots, a_r >$ be a quadratic form over a field K (of characteristic not 2) and $L = K(\sqrt{d})$ be a quadratic extension of K. If q is not isotropic over K and is isotropic over L, then q is equivalent to $\lambda < 1, -d, b_1, b_2, \cdots, b_{r-2} >$, with $\lambda, b_1, \cdots, b_{r-2} \in K$.*

Proof. Let $\lambda_i + \mu_i\sqrt{d} \in L$, for $1 \leq i \leq r$, such that $\sum_{1\leq i\leq r} a_i(\lambda_i + \mu_i\sqrt{d})^2 = 0$, so that $\sum_{1\leq i\leq r} a_i{\lambda_i}^2 + d\sum_{1\leq i\leq r} a_i{\mu_i}^2 = 0$ and $\sum_{1\leq i\leq r} a_i\lambda_i\mu_i = 0$. Since q is not isotropic over K, $\sum_{1\leq i\leq r} a_i{\lambda_i}^2 \neq 0$, $\sum_{1\leq i\leq r} a_i{\mu_i}^2 \neq 0$ and $d = -\frac{\sum_{1\leq i\leq r} a_i{\lambda_i}^2}{\sum_{1\leq i\leq r} a_i{\mu_i}^2} = \frac{-q(\bar{\lambda})}{q(\bar{\mu})}$, where $\bar{\lambda} = (\lambda_1, \cdots, \lambda_r)$ and $\bar{\mu} = (\mu_1, \cdots, \mu_r)$. The equation $\sum_{1\leq i\leq r} a_i\lambda_i\mu_i = 0$ shows first that $\bar{\lambda}$ and $\bar{\mu}$ are linearly independent vectors in K^r and that if we extend $\bar{\lambda}, \bar{\mu}$, to a suitable basis of K^r, then q has the form $< -dq(\mu), q(\mu), b_1, \cdots, b_{r-2} >$. This proves the lemma.

Let $L = K(\sqrt{d})$. Let w be a valuation of L, which extends a valuation v of K and let L_w and K_v be their respective completions. Then L_w contains $K_v(\sqrt{d})$. Since f is equivalent to $< 1, -a, -b, ab >$ over L and f is isotropic over K_v, f is isotropic over L_w. Therefore by our previous consideration, f is isotropic over L. By the above lemma, f is equivalent to $< \lambda, -\lambda d, \mu, \mu' >$. Therefore comparing discriminants, we have $d = abc = -\lambda^2 d\mu\mu'$, so that f is equivalent to $< \lambda, -\lambda d, \mu, -\mu >$. Hence obviously, f is isotropic, as it contains $< \mu, -\mu >$.

Let $r \geq 5$ and $f = < a_1, a_2, a_3, a_4, a_5, \cdots, a_r >$, $a_i \in K$, for $1 \leq i \leq r$. Let S be a finite set of valuations of K such that S contains the 2-adic valuations i.e., the valuations of K lying over the prime 2 of $\mathbb{Q}$, the archimedean valuations and such that a_3, a_4, a_5 are units in K_v for $v \notin S$. We can choose such a finite set, since for any element $a \in K$, $v(a) \leq 1$ for all but a finite number of valuations v of K and if $a \neq 0$, applying the above remark to a^{-1} too, we have that $v(a) = 1$ for all but a finite number of valuations v of K, i.e., a is a unit in K_v for all but a finite set of valuations v of K. By Proposition 5, it follows that the quadratic form $< a_3, a_4, a_5 >$ and hence $< a_3, a_4, a_5, \cdots, a_r >$ is isotropic over K_v for $v \notin S$. We now claim that for any $v \in S$, there exists a $\mu_v \in K_v^*$, which is a value of $< a_1, a_2 >$ and such that $-\mu_v$ is a value of $< a_3, a_4, a_5, \cdots, a_r >$. To prove this claim, we consider two cases.

Case.1: Suppose $< a_1, a_2 >$ is anisotropic over K_v. Since f is isotropic over K_v, there exists $(\lambda_1, \lambda_2, \cdots, \lambda_r) \in K_v^r$, such that $\sum_{1 \leq i \leq r} a_i \lambda_i^2 = 0$. We then choose μ_v to be $a_1\lambda_1^2 + a_2\lambda_2^2$, which obviously cannot be zero.

Case.2: Suppose $< a_1, a_2 >$ is isotropic over K_v. Then choose μ_v, to be any non-zero element represented by $< a_3, a_4, \cdots, a_r >$. Since $< a_1, a_2 >$ is isotropic over K_v, it represents all elements of K_v, in particular μ_v.

By Lemma 1, there exist $x, y \in K$ such that x is close to x_v and y is close to y_v for $v \in S$, so that $\mu = a_1x^2 + a_2y^2$ is close to μ_v and in fact belongs to the same square class as μ_v for every $v \in S$. (If v is not archimedean and $\mu\mu_v^{-1} \equiv 1 mod(\pi_v)$, then by an argument similar to the one employed in the proof of proposition 5, $\mu\mu_v^{-1}$ is a square. If v is archimedean and real, this simply means that μ and μ_v have the same sign). Thus the form $< a_1, a_2 >$ represents μ over K, so that $< a_1, a_2 >$ is equivalent to $< \lambda, \mu >$ for $\lambda \in K$ and μ and μ_v are in the same square class for $v \in S$. Since for $v \in S$, $< a_3, a_4, \cdots, a_r >$ represents $-\mu_v$, it also represents $-\mu$ over K_v (since these elements have the same square class).

Thus the form $< \mu, a_3, a_4, \cdots, a_r >$ is isotropic over K_v for $v \in S$. For $v \notin S$, by Proposition 5, $< a_3, a_4, a_5 >$ and hence $< a_3, a_4, \cdots, a_r >$ is isotropic over K_v and a fortiori, $< \mu, a_3, a_4, \cdots, a_r >$ is isotropic over K_v, so that $< \mu, a_3, \cdots, a_r >$ is isotropic over K_v, for all v. By induction on r, it follows that $< \mu, a_3, \cdots, a_r >$ is isotropic over K, so that $< a_3, a_4, \cdots, a_r >$ represents $-\mu$ over K and $f =< a_1, a_2, a_3, \cdots, a_r >=< \lambda, \mu, -\mu, \cdots >$. Hence f contains $\mu < 1, -1 >$ and therefore it is isotropic, which proves the theorem.

R. Sridharan
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: sridhar@math.tifr.res.in