5-2006

The Hasse-Minkowski Theorem

Adam Gamzon
University of Connecticut

The Hasse–Minkowski Theorem

by
Adam Gamzon

a thesis
presented to the Department of Mathematics
in partial fulfillment
of the requirements for the degree
of Bachelor of Arts with Honors

University of Connecticut
Storrs, Connecticut
March 8, 2006

In memory of my grandfather Herman Sholovitz

Contents

Introduction

The Hasse–Minkowski theorem concerns the classification of quadratic forms over global fields (i.e., finite extensions of either $\mathbb{Q}$ or rational function fields with a finite constant field). Hasse proved the theorem over the rational numbers in his Ph.D. thesis in 1921. He extended the research of his thesis to quadratic forms over all number fields in 1924. Then Rauter, a Ph.D. student of Hasse's, proved the theorem over $\mathbb{F}_p(T)$ (where $p > 2$) for his Ph.D. thesis. However, there is really no change that needs to be made to do the same with a general finite constant field of odd characteristic. Historically, the Hasse–Minkowski theorem was the first notable application of p-adic fields that caught the attention of a wide mathematical audience. The goal of this thesis is to discuss the Hasse–Minkowski theorem over the rational numbers and over the rational function fields with a finite constant field of odd characteristic. Our treatments of quadratic forms and local fields, though, are more general than what is strictly necessary for our proofs of the Hasse–Minkowski theorem over $\mathbb{Q}$ and its analogue over rational function fields (of odd characteristic). Not only should these topics be studied for their own sake, but, as mentioned before, one can also contemplate the Hasse–Minkowski theorem in a more general context than what this thesis will cover. We hope that this thesis can be useful as a reference for other students.

Throughout the following we will assume the reader to be familiar with linear algebra, number theory, finite fields, finite field extensions, group theory, and point-set topology (in metric spaces). In addition, the reader should have at least seen p-adic fields in some context beforehand. All of our vector spaces are assumed to be finite-dimensional and $\mathbb{F}$ denotes a finite field of odd characteristic unless otherwise stated (e.g., in Section 3.1).

Chapter 1 gives some background information on the Hasse–Minkowski theorem while opening up some interesting questions to be answered later in the thesis. In Chapter 2 we will discuss the purely algebraic aspects of quadratic forms over a field, including orthogonality, equivalence, and Witt cancellation. We conclude Chapter 2 with a classification of non-degenerate quadratic forms over $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{F}$. In Chapter 3 the reader will be introduced to local fields and quadratic forms over local fields using the completions of $\mathbb{Q}$ and $\mathbb{F}(T)$ as specific examples. We will also introduce the Hilbert symbol and use it to classify non-degenerate quadratic forms over local fields not of characteristic two. Chapter 4 covers the proof of the Hasse–Minkowski theorem over $\mathbb{Q}$ and some applications while Chapter 5 covers an analogous

proof of the Hasse–Minkowski theorem over $\mathbb{F}(T)$.

A special thanks to Keith Conrad for his infinite patience, guidance and wisdom. Without his help none of this would have been written. Also, thank you to my friends and family, especially my parents and my fiancé Allison Munk, for bearing with me throughout the long hours that I stayed locked in my room.

Chapter 1

Overview

A quadratic form in a concrete sense is a homogeneous degree two polynomial. In other words, a quadratic form (over a commutative ring A) is a function $Q : A^n \to A$ defined by

$$Q(x_1, \ldots, x_n) = \sum_{1 \leq i \leq j \leq n} a_{ij} x_i x_j$$

where $a_{ij} \in A$. The simplest example of a quadratic form is

$$x_1^2 + \cdots + x_n^2.$$

Another example (with $n = 2$) is $x^2 + xy + y^2$. Similar to the way studying linear algebra over fields (vector spaces) is simpler than studying it over rings (modules), quadratic forms are easier to analyze over fields. However, number-theorists were historically most interested in quadratic forms over $\mathbb{Z}$. The representation problem for a quadratic form $Q : A^n \to A$ asks, for $a \in A$, if there are $a_1, \ldots, a_n$ in A such that $Q(a_1, \ldots, a_n) = a$. The Hasse–Minkowski theorem answers this question for $A = \mathbb{Q}$ and for $A = \mathbb{F}(T)$.

When $A = F$ is a field not of characteristic 2, any quadratic form over F can be diagonalized, that is, have its cross terms with i not equal to j removed, by a linear change of variables (we will see why later). For example, the quadratic form $x^2 + xy + y^2$ over $\mathbb{Q}$ can be written as $x'^2 + 3y'^2$ where $x' = x + \frac{1}{2}y$ and $y' = \frac{1}{2}y$. So our study of quadratic forms over $\mathbb{Q}$ and $\mathbb{F}(T)$ can focus on diagonal quadratic forms with no essential loss of generality.

1.1 Examples

The following four theorems are famous results about integers represented by quadratic forms over $\mathbb{Z}$. Elementary number theory courses may prove the first two theorems, but the second two are more advanced.

Theorem 1.1 (Fermat)**.** *A prime p has the form $x^2 + y^2$ with $x, y \in \mathbb{Z}$ if and only if $p \equiv 1 \bmod 4$ or $p = 2$.*

Theorem 1.2 (Fermat)**.** *A prime p has the form $x^2 - 2y^2$ with $x, y \in \mathbb{Z}$ if and only if $p \equiv \pm 1 \bmod 8$ or $p = 2$.*

Theorem 1.3 (Legendre)**.** *Write a positive integer n in the form $4^a n'$ with $a \geq 0$ and $n' \not\equiv 0 \bmod 4$. Then n is a sum of three integer squares if and only if $n' \not\equiv 7 \bmod 8$.*

Theorem 1.4 (Lagrange)**.** *Every positive integer is a sum of four integer squares.*

Quadratic forms over $\mathbb{Q}$ are sometimes sufficient to answer questions about quadratic forms over $\mathbb{Z}$. Here are two results along these lines.

Theorem 1.5. *If an integer is a sum of two rational squares then it is a sum of two integer squares.*

Theorem 1.6. *If an integer is a sum of three rational squares then it is a sum of three integer squares.*

We will use Theorem 1.6 to reduce the proof of Legendre's theorem to a question of an integer being represented as a sum of three rational squares, which will be answered using the Hasse–Minkowski theorem for $x^2 + y^2 + z^2$. The reader can find proofs of Theorems 1.5 and 1.6 in Appendix A. We will prove Theorems 1.3 and 1.4 in Section 4.5.

In addition to the representation problem, another question one can ask about quadratic forms is: what are all the different ways that a quadratic form takes on a value? For example, classifying Pythagorean triples is the same thing as finding all rational points on the unit circle (if $a^2 + b^2 = c^2$ then $(a/c)^2 + (b/c)^2 = 1$).

Definition 1.7. A *plane conic* is a level curve of a two-variable quadratic form Q over a field F. In other words, it is the set

$$\{(x, y) \in F^2 : Q(x, y) = d\}$$

for some $d \in F$.

For "nice" plane conics over $\mathbb{Q}$, knowing one rational solution allows us to write down an explicit parameterization for every other rational solution in terms of the initial solution and a rational parameter.

Example 1.8. Let the plane conic be the circle $x^2 + y^2 = 1$. Using the rational point $(-1, 0)$ and some geometry we can get a formula for every other rational point on the unit circle:

$$\left(\frac{-m^2+1}{m^2+1}, \frac{2m}{m^2+1}\right)$$

for $m \in \mathbb{Q}$. For $m = \infty$ we get the original point $(-1, 0)$ back.

Example 1.9. Let the plane conic be the circle $x^2 + y^2 = 2$. One rational point on this circle is $(1, 1)$ and by a geometric argument with this point we can get a formula for every other rational point on this circle:

$$\left(\frac{m^2-2m-1}{m^2+1}, \frac{-m^2-2m+1}{m^2+1}\right)$$

for $m \in \mathbb{Q}$ with the exception of $(1, -1)$ which corresponds to $m = \infty$. The original point occurs at $m = -1$.

Example 1.10. The plane conic $x^2 + y^2 = 3$ has many real points but no rational points. In other words, there are no $x, y \in \mathbb{Q}$ that satisfy $x^2+y^2 = 3$. If there were then by Theorem 1.5 there would be $x, y \in \mathbb{Z}$ that satisfy $x^2 + y^2 = 3$. This is false, however, since 3 is not the sum of two integer squares.

The general case of Examples 1.8 and 1.9 is the next result.

Theorem 1.11. *Suppose there is a rational solution (x_0, y_0) to the equation $ax^2 + by^2 = c$ for non-zero $a, b, c \in \mathbb{Q}$. Then every other rational solution is of the form*

$$\left(\frac{bm^2x_0 - 2bmy_0 - ax_0}{bm^2+a}, \frac{-bm^2y_0 - 2amx_0 + ay_0}{bm^2+a}\right)$$

for $m \in \mathbb{Q}$ with $bm^2 + a \neq 0$ or $m = \infty$ (namely $(x_0, -y_0)$).

Including ∞ may seem weird, so a reformulation and proof of Theorem 1.11 that clears up the business of $m = \infty$ is in Appendix A. Thus the task of finding all the rational points on a diagonal plane conic over $\mathbb{Q}$ boils down to knowing the existence of one rational point. The Hasse–Minkowski theorem will let us decide when there is a rational point.

Quadratic forms also arise in other contexts as "algebraic invariants". For instance, we can attach a quadratic form to any finite field extension. If F is a field and K is a finite extension of F, then we define a quadratic form Q over F, called the *trace form*, by $Q(\alpha) = \mathrm{Tr}_{K/F}(\alpha^2)$ for $\alpha \in K$. Picking an F-basis of K allows us to express Q as a specific homogeneous polynomial.

Example 1.12. Let $F = \mathbb{R}$ and $K = \mathbb{C}$. Using the basis $\{1, i\}$, write $z \in \mathbb{C}$ as $z = x + iy$ with $x, y \in \mathbb{R}$. Then

$$Q(z) = \mathrm{Tr}_{\mathbb{C}/\mathbb{R}}(z^2) = 2x^2 - 2y^2.$$

Example 1.13. Let $K = \mathbb{Q}(\theta)$, where θ is a root of $T^4 - 8T + 9$. This is irreducible over $\mathbb{Q}$ because it is irreducible modulo 5. View K as a 4-dimensional $\mathbb{Q}$-vector space. In the basis $\{1, \theta, \theta^2, \theta^3\}$, the multiplication-by-$\theta$ map has matrix

$$[m_\theta] = \begin{pmatrix} 0 & 0 & 0 & -9 \\ 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

So for $\alpha \in K$, writing $\alpha = x + y\theta + z\theta^2 + t\theta^3$ for $x, y, z, t \in \mathbb{Q}$, the trace form of K over $\mathbb{Q}$ is

$$\begin{aligned} Q(\alpha) &= \mathrm{Tr}_{K/\mathbb{Q}}(\alpha^2) \\ &= \mathrm{Tr}_{K/\mathbb{Q}}([m_\alpha]^2) \\ &= 4x^2 - 36z^2 + 192t^2 + 48xt + 48yz - 72yt. \end{aligned} \tag{1.1}$$

This is a quadratic form over $\mathbb{Q}$. Notice it is not in diagonal form.

Example 1.14. Similarly, the polynomial $T^4 + 6T^2 - 4T + 6$ is irreducible over $\mathbb{Q}$ because it is irreducible modulo 5. Let $L = \mathbb{Q}(\theta')$ where θ' is a root of this polynomial. Viewing L as a $\mathbb{Q}$-vector space, we have that the multiplication-by-θ' map has matrix

$$[m_{\theta'}] = \begin{pmatrix} 0 & 0 & 0 & -6 \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -6 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

in the basis $\{1, \theta', \theta'^2, \theta'^3\}$. Write $\alpha \in L$ as $\alpha = x + y\theta' + z\theta'^2 + t\theta'^3$ for $x, y, z, t \in \mathbb{Q}$. Then the trace form of L over $\mathbb{Q}$ is

$$\begin{aligned} Q(\alpha) &= \mathrm{Tr}_{L/\mathbb{Q}}(\alpha^2) \\ &= \mathrm{Tr}_{L/\mathbb{Q}}([m_\alpha]^2) \\ &= 4x^2 - 12y^2 + 48z^2 - 168t^2 - 24xz + 24xt + 24yz + 96yt - 240zt, \end{aligned}$$

which is a (non-diagonal) quadratic form over $\mathbb{Q}$.

In Section 4.5 we will show that the fields in Examples 1.13 and 1.14 are not isomorphic by comparing their trace forms.

Remark 1.15. If K/F is inseparable, then the trace map $\mathrm{Tr}_{K/F}$ is identically zero so the trace form is also identically zero.

1.2 The Statement of the Hasse–Minkowski Theorem

In general, it is not as easy as in Example 1.8 to decide if other rational quadratic forms represent a given rational number over $\mathbb{Q}$. One can see this difficulty already with the quadratic form in Example 1.13. For instance, does it take on the value 3 with $x, y, z, t \in \mathbb{Q}$? We will answer this question in Example 2.55. Hasse, at the suggestion of Hensel, made the study of such problems the main topic of his doctoral dissertation which he published in 1923. His research led him to reduce the question of representations over $\mathbb{Q}$ to knowing if a rational quadratic form represents a rational number over every metric completion of $\mathbb{Q}$.

In addition to the usual absolute value on $\mathbb{Q}$, for each prime p there is another absolute value on $\mathbb{Q}$ which is called the *p-adic absolute value*. Completing $\mathbb{Q}$ with respect to a p-adic absolute value yields a complete field called a p-adic field, which is denoted by $\mathbb{Q}_p$. Within number theory, the p-adic fields are as significant as the complete field $\mathbb{R}$.

Theorem 1.16 (Hasse–Minkowski)**.** *Let $Q(x_1, \ldots, x_n)$ be a quadratic form over $\mathbb{Q}$ in any dimension $n \geq 1$.*

1) *Given $r \in \mathbb{Q}^\times$, the equation $Q(x_1, \ldots, x_n) = r$ is solvable over $\mathbb{Q}$ if and only if it is solvable over $\mathbb{R}$ and every $\mathbb{Q}_p$.*

2) *The equation* $Q(x_1, \ldots, x_n) = 0$ *is non-trivially solvable over* $\mathbb{Q}$ *if and only if it is non-trivially solvable over* $\mathbb{R}$ *and every* $\mathbb{Q}_p$.

Remark 1.17. In the Hasse–Minkowski theorem, non-trivially solvable means that there is a solution such that not all x_i's are zero.

A corollary to the Hasse–Minkowski theorem states that two quadratic forms are equal after a linear change of variables over $\mathbb{Q}$ if and only if they are so over $\mathbb{R}$ and every $\mathbb{Q}_p$. This is referred to as the weak Hasse–Minkowski theorem.

An analogue of the Hasse–Minkowski theorem also holds true for quadratic forms over the rational function field $\mathbb{F}(T)$ where $\mathbb{F}$ is a finite field. There is a completion of $\mathbb{F}(T)$ with respect to an absolute value corresponding to each monic irreducible polynomial π of $\mathbb{F}[T]$. In addition, there is another absolute value that induces a completion of $\mathbb{F}(T)$ that is in some ways superficially like $\mathbb{R}$ (namely $\mathbb{F}[T]$ is discrete in it as $\mathbb{Z}$ is discrete in $\mathbb{R}$), but is in most ways like the p-adic fields. Using these completions we will prove the Hasse–Minkowski theorem over $\mathbb{F}(T)$.

This idea of looking for information "locally" (i.e., in every completion of $\mathbb{Q}$ or $\mathbb{F}(T)$) to determine the answer to a question "globally" (i.e., over $\mathbb{Q}$ or $\mathbb{F}(T)$) is known as the local-global principle. Although this principle plays an influential role in number theory, it is not true in general.

Example 1.18 (Selmer)**.** As a counterexample to the local-global principle in higher degree, the equation $3x^3 + 4y^3 + 5z^3 = 0$ has a non-trivial solution (i.e., other than (0,0,0)) over $\mathbb{R}$ and all $\mathbb{Q}_p$, but it does not have a non-trivial solution over $\mathbb{Q}$. See [1, p. 72] for a further discussion of this and other counterexamples.

The Hasse–Minkowski theorem might seem to be creating more difficulties in deciding if a rational quadratic form represents a rational number over $\mathbb{Q}$ because there are an infinite number of completions. It turns out, however, that not only is there just a finite number of primes p in any particular example that one needs to check for representation over $\mathbb{Q}_p$, but also there is a relatively easy way to check the representation problem in these fields (and $\mathbb{R}$).

Chapter 2

Algebraic Properties of Quadratic Forms

The goal of this chapter is to move from the concrete description of a quadratic form as a homogeneous degree two polynomial to a coordinate-free description as a "quadratic function" on a vector space. Furthermore, all of the concepts about quadratic forms that are necessary for the proof of the Hasse-Minkowksi theorem that do not rely upon number theory will be covered here.

Consider the relationship between the standard inner product on $\mathbb{R}^n$ and a sum of n squares, $Q(\mathbf{x}) = x_1^2 + x_2^2 + \cdots + x_n^2$. This can be interpreted as squared length in $\mathbb{R}^n$. In other words, $Q(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = ||\mathbf{x}||^2$. Furthermore, the inner product can be expressed in terms of Q:

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2}(Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})).$$

We will extend the above relationship between sums of squares and the inner product on $\mathbb{R}^n$ to a coordinate-free definition of a quadratic form in terms of a symmetric bilinear form on a vector space. Throughout this chapter V is a (finite-dimensional) vector space over a field F.

2.1 Bilinear Forms

Definition 2.1. A function $\mathrm{B} : V \times V \to F$ is called a *bilinear form* on V if B is linear in each variable when the other variable is fixed. In other words,

1) $\mathrm{B}(c_1v_1 + c_2v_2, w) = c_1\mathrm{B}(v_1, w) + c_2\mathrm{B}(v_2, w)$ for $c_i \in F$ and $v_i, w \in V$,

and

2) $\mathrm{B}(v, c_1w_1 + c_2w_2) = c_1\mathrm{B}(v, w_1) + c_2\mathrm{B}(v, w_2)$ for $c_i \in F$ and $v, w_i \in V$.

When $n = \dim_F V$, choosing a basis $\{e_1, \ldots, e_n\}$ for V allows us to express B concretely:

$$\mathrm{B}(v, w) = \mathrm{B}\left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i,j=1}^{n} a_{ij} x_i y_j,$$

where $a_{ij} = \mathrm{B}(e_i, e_j) \in F$. The matrix $A = (a_{ij}) = (\mathrm{B}(e_i, e_j))$ lets us give another formula for B in terms of the "column vector" isomorphism $[\,] : V \to F^n$ associated to the basis $\{e_1, \ldots, e_n\}$, namely: $\mathrm{B}(v, w) = [v] \cdot A\,[w]$, where $\cdot$ is the inner product on F^n. We call A the *matrix representation* of B in the basis $\{e_1, \ldots, e_n\}$. Note that in general $[v] \cdot A[w] = A^\top[v] \cdot [w]$, where $A^\top$ is the transpose of A. This allows us to flip between using matrices "on the left" and using matrices "on the right" to express B. We will use them on the right.

Additionally, since A is determined by the way B acts on basis vectors, for a fixed basis $\{e_1, \ldots, e_n\}$ we get a bijection from bilinear forms on V to $\mathrm{M}_n(F)$ by sending B to the matrix $(\mathrm{B}(e_i, e_j))$. This is one-to-one since if B_1 and B_2 are bilinear forms with the same matrix representation, then B_1 and B_2 act on the basis vectors in exactly the same way (i.e., $\mathrm{B}_1(e_i, e_j) = \mathrm{B}_2(e_i, e_j)$), so B_1 and B_2 act on all vectors in V in the same way since every vector is a linear combination of $e_1, \ldots, e_n$. The surjectivity is clear because given any matrix $A = (a_{ij})$ we can define a bilinear form on V by $\mathrm{B}(v, w) = [v] \cdot A[w]$ and the matrix representation of B is A.

Definition 2.2. A bilinear form B on V is called *symmetric* if

$$\mathrm{B}(v, w) = \mathrm{B}(w, v)$$

for all $v, w \in V$.

In terms of a basis $\{e_1, \ldots, e_n\}$ of V, B is symmetric if and only if $\mathrm{B}(e_i, e_j) = \mathrm{B}(e_j, e_i)$ for all i, j. So B is symmetric if and only if any matrix representation is a symmetric matrix.

Example 2.3. The inner product on $\mathbb{R}^n$,

$$\mathrm{B}(v, w) = v \cdot w = x_1 y_1 + \cdots + x_n y_n,$$

is a symmetric bilinear form.

Example 2.4. On $\mathbb{R}^2$, the function

$$\begin{aligned} \mathrm{B}((x_1, y_1), (x_2, y_2)) &= x_1 x_2 + y_1 x_2 + y_1 y_2 \\ &= \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \end{aligned}$$

is bilinear, but not symmetric.

Definition 2.5. A bilinear form B is called *non-degenerate* if for any $v \in V$ where $v \neq 0$, $\mathrm{B}(v, w) \neq 0$ for some $w \in V$. A bilinear form is called *degenerate* if it is not non-degenerate.

Theorem 2.6. *A bilinear form is non-degenerate if and only if any matrix representation of the bilinear form is invertible.*

Proof. Choose any basis $\{e_1, \ldots, e_n\}$ for V. Let B be a bilinear form on V and let A be the matrix representation of B in the chosen basis.

Suppose A is invertible. For any non-zero $v \in V$ write

$$[v] = [c_1 e_1 + \cdots + c_n e_n] = (c_1, \ldots, c_n)$$

where some $c_i \neq 0$. Since A is surjective on F^n, take w_0 such that $A[w_0] = [e_i]$. Thus,

$$\begin{aligned} \mathrm{B}(v, w_0) &= [v] \cdot A[w_0] \\ &= (c_1, \ldots, c_i, \ldots, c_n) \cdot (0, \ldots, 0, 1, 0, \ldots, 0) \\ &= c_i \\ &\neq 0. \end{aligned}$$

Hence B is non-degenerate.

We now show the reverse implication by proving the contrapositive. Suppose that A is not invertible. Since $\det A^\top = \det A$, $A^\top$ is also not invertible. Then some $v \in V$, where $v \neq 0$, satisfies $A^\top[v] = 0$. So for all $w \in V$,

$$\begin{aligned} \mathrm{B}(v, w) &= [v] \cdot A[w] \\ &= A^\top[v] \cdot [w] \\ &= 0 \cdot [w] \\ &= 0. \end{aligned}$$

Thus, B is degenerate. □

Example 2.7. On F^3, the symmetric bilinear form

$$\mathrm{B}(v,w) = v \cdot \begin{pmatrix} 2 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix} w$$

is non-degenerate since the matrix has determinant $-1 \neq 0$.

Fixing $v \in V$, we can contemplate the function $V \to F$ by $w \mapsto \mathrm{B}(v,w)$. This function, denoted $\mathrm{B}(v,-)$, lies in the dual space V^* since B is linear in its second component. Additionally, the function $V \to V^*$ by $v \mapsto \mathrm{B}(v,-)$ is linear because B is linear in its first component.

Theorem 2.8. *Let* B *be a bilinear form on* V*. The following are equivalent:*

1) B *is non-degenerate,*

2) *the map* $f : V \to V^*$ *defined by* $v \mapsto \mathrm{B}(v,-)$ *is an isomorphism.*

Proof. Let B be non-degenerate. Since f is linear and $\dim V = \dim V^*$ it is enough to show that f is one-to-one. Suppose $v \in \ker(f)$, so $\mathrm{B}(v,w) = 0$ for all $w \in V$. Then $v = 0$ by non-degeneracy. Therefore f is an isomorphism.

On the other hand, let f be an isomorphism. For $v \in V$, $v \neq 0$, we have $f(v) \neq 0$ so $\mathrm{B}(v,w) \neq 0$ for some $w \in V$. Hence B is non-degenerate. □

Remark 2.9. Given any linear map $f : V \to V^*$, there is a bilinear form $\mathrm{B} : V \times V \to F$ defined by $\mathrm{B}(v,w) = f(v)(w)$. So a bilinear form on V is a particular choice of linear map $V \to V^*$ and a non-degenerate bilinear form is a particular choice of isomorphism $V \to V^*$.

For a bilinear form B on V and a subspace $W \subset V$ we have an induced bilinear form $\mathrm{B}|_W$ by restricting B to W.

Definition 2.10. Let B be a bilinear form on V. A subspace $W \subset V$ is called *non-degenerate* when $\mathrm{B}|_W$ is non-degenerate. Otherwise, we call W *degenerate*.

Example 2.11. In Example 2.7, F^3 is non-degenerate relative to B but the line $F(0,1,1)$ is a degenerate subspace since $\mathrm{B}((0,a,a),(0,b,b)) = 0$ for all $a, b \in F$.

Definition 2.12. The bilinear forms B_1 on V_1 and B_2 on V_2 are *equivalent* if there is a linear isomorphism $C : V_1 \to V_2$ such that $\mathrm{B}_2(Cv, Cw) = \mathrm{B}_1(v, w)$ for all $v, w \in V_1$. Denote this by $\mathrm{B}_1 \cong \mathrm{B}_2$.

Equivalence of bilinear forms is clearly an equivalence relation.

Theorem 2.13. *Let* B_1 *and* B_2 *be bilinear forms on* V_1 *and* V_2*, where* $\dim V_1 = \dim V_2$*, and let* A_1 *and* A_2 *be matrix representations of* B_1 *and* B_2 *in some bases of* V_1 *and* V_2*. Then* B_1 *and* B_2 *are equivalent if and only if* $A_1 = M^\top A_2 M$ *for some* $M \in \mathrm{GL}_n(F)$*, where* $n = \dim V_i$*.*

Proof. Let $[\ ]_1 : V_1 \to F^n$ be the column isomorphism of the basis in which A_1 represents B_1. Let $[\ ]_2 : V_2 \to F^n$ be the column isomorphism of the basis in which A_2 represents B_2.

Suppose $\mathrm{B}_1 \cong \mathrm{B}_2$, using a linear isomorphism C as in Definition 2.12. Then $\mathrm{B}_2(Cv, Cw) = \mathrm{B}_1(v, w)$ for all $v, w \in V_1$. Choose $M : F^n \to F^n$ such that it makes the following diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{[\ ]_1} & F^n \\ \big\downarrow{\scriptstyle C} & & \big\downarrow{\scriptstyle M} \\ V_2 & \xrightarrow{[\ ]_2} & F^n \end{array}$$

commute. Clearly $M \in \mathrm{GL}_n(F)$ and $M[v]_1 = [Cv]_2$. Then for all v, w in V_1,

$$\begin{aligned} \mathrm{B}_1(v, w) &= \mathrm{B}_2(Cv, Cw) \\ [v]_1 \cdot A_1 [w]_1 &= [Cv]_2 \cdot A_2 [Cw]_2 \\ &= M[v]_1 \cdot A_2 M[w]_1 \\ &= [v]_1 \cdot M^\top A_2 M[w]_1, \end{aligned}$$

so $A_1 = M^\top A_2 M$.

Conversely, suppose $A_1 = M^\top A_2 M$ for some $M \in \mathrm{GL}_n(F)$. Choose C such that the above diagram commutes. Then C is a linear isomorphism from V_1 to V_2 and again $M[v]_1 = [Cv]_2$. Reversing the calculations,

$$\begin{aligned} \mathrm{B}_1(v, w) &= [v]_1 \cdot A_1 [w]_1 \\ &= [v]_1 \cdot M^\top A_2 M[w]_1 \\ &= M[v]_1 \cdot A_2 M[w]_1 \\ &= [Cv]_2 \cdot A_2 [Cw]_2 \\ &= \mathrm{B}_2(Cv, Cw), \end{aligned}$$

so $\mathrm{B}_1 \cong \mathrm{B}_2$. □

Corollary 2.14. *If A_1 and A_2 are matrix representations of a bilinear form* B *in two different bases of V, then $A_1 = M^\top A_2 M$ for some $M \in \mathrm{GL}_n(F)$ where $n = \dim V$.*

Proof. Set $\mathrm{B}_1 = \mathrm{B}_2 = \mathrm{B}$ and $V_1 = V_2 = V$ in Theorem 2.13. □

Definition 2.15. Let B be a bilinear form on V. Its *discriminant* is

$$\mathrm{disc}(\mathrm{B}) = \det(\mathrm{B}(e_i, e_j))$$

where $\{e_1, \ldots, e_n\}$ is any basis of V.

If we change bases, how does the discriminant of B change? If A_1 and A_2 are matrices representing B, then $A_1 = M^\top A_2 M$ for some $M \in \mathrm{GL}_n(F)$, so $\det A_1 = (\det A_2)(\det M)^2$. Hence the discriminant is well-defined up to non-zero squares: $\mathrm{disc}(\mathrm{B}) \in F^\times / F^{\times 2}$ or $\mathrm{disc}(\mathrm{B}) = 0$. For $x, y \in F$, if $x = yz^2$ for some $z \in F^\times$ then we will denote this by $x \sim y$. This is an equivalence relation on F.

Definition 2.16. For a bilinear form B on V, we say that vectors $v, w \in V$ are *orthogonal* (with respect to B) if $\mathrm{B}(v, w) = 0$. This is denoted by $v \perp w$. If $v \perp u$ for all u in a subspace $U \subset V$ then we write $v \perp U$.

Definition 2.17. Let $n \geq 2$. A basis $\{e_1, \ldots, e_n\}$ of V is called *orthogonal* (with respect to B) if $e_i \perp e_j$ (i.e., $\mathrm{B}(e_i, e_j) = 0$) for all $i \neq j$.

Remark 2.18. When $\dim V = 1$, any non-zero vector is considered to be an orthogonal basis for V.

The orthogonality relation behaves linearly: if $v \perp w_1$ and $v \perp w_2$ then

$$v \perp (c_1 w_1 + c_2 w_2)$$

for $c_1, c_2 \in F$. If V has an orthogonal basis, then the matrix $(\mathrm{B}(e_i, e_j))$ is diagonal so it is symmetric. This means B is a symmetric bilinear form. In particular, V can only have an orthogonal basis when B is symmetric. For this reason, throughout the remainder of this section B *will be a symmetric bilinear form.*

Remark 2.19. The orthogonality relation is not necessarily symmetric when B is not symmetric. To see this, consider the bilinear form in Example 2.4. We have $(1,0) \perp (0,1)$, but $(0,1) \not\perp (1,0)$.

In addition to orthogonal vectors, there are also subspaces called *orthogonal*: for subspaces U and W in V, write $U \perp W$ when $u \perp w$ for all $u \in U$ and $w \in W$. For a subspace $U \subset V$, its orthogonal space is $U^\perp = \{v \in V : v \perp U\}$. So $U^\perp \perp U$. We abbreviate $(Fv)^\perp$ to $v^\perp$. Note that $U \cap U^\perp = \{0\}$ exactly when U is a non-degenerate subspace. Also note $U \cap U^\perp$ can be non-zero: if $\mathrm{B}(v,v) = 0$ for some $v \neq 0$ and we use $U = Fv$, then $U \subset U^\perp$.

If B is non-degenerate (and symmetric) on V, then for any subspace U there is an exact sequence:

$$0 \to U^\perp \to V \to U^* \to 0$$

where $U^\perp \to V$ is the natural inclusion and the map $V \to U^*$ is defined by $v \mapsto \mathrm{B}(v,-)|_U$ (with kernel $U^\perp$). So

$$\dim U + \dim U^\perp = \dim V. \tag{2.1}$$

It then follows that $(U^\perp)^\perp = U$ for all $U \subset V$:

$$\dim V = \dim U + \dim U^\perp = \dim U^\perp + \dim (U^\perp)^\perp,$$

so $\dim U = \dim (U^\perp)^\perp$ and $U \subset (U^\perp)^\perp$. Hence $U_1 \subset U_2 \iff U_2^\perp \subset U_1^\perp$.

Lemma 2.20. *If F does not have characteristic 2 and $\mathrm{B}(v,v) = 0$ for all $v \in V$ then $\mathrm{B}(v,w) = 0$ for all $v, w \in V$.*

Proof. We have

$$\begin{aligned} \mathrm{B}(v,w) &= \frac{1}{2}(\mathrm{B}(v+w, v+w) - \mathrm{B}(v,v) - \mathrm{B}(w,w)) \\ &= 0. \end{aligned}$$

□

Remark 2.21. Lemma 2.20 is false when F has characteristic 2; the symmetric bilinear form $\mathrm{B}(v,w) = v \cdot \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) w$ on F^2 has $\mathrm{B}(v,v) = 0$ for all v while $\mathrm{B} \not\equiv 0$.

Theorem 2.22. *Let* B *be a symmetric bilinear form on* V *with* $\dim V > 0$ *and* $\mathrm{char}(F) \neq 2$. *Then*

1) V *admits an orthogonal basis,*

2) *any orthogonal basis of a non-degenerate subspace* $U \subset V$ *can be extended to an orthogonal basis of* V *and* $V = U \oplus U^{\perp}$.

Proof. 1) Let $n = \dim V$. We will proceed by induction on n. For $n = 1$, the statement is vacuously true. Now suppose $n \geq 2$. If $\mathrm{B} \equiv 0$ on V then any basis of V is orthogonal. Otherwise $\mathrm{B} \not\equiv 0$, so by Lemma 2.20 we can pick $v_1 \in V$ such that $\mathrm{B}(v_1, v_1) \neq 0$. Consider the map $\mathrm{B}(v_1, -) : V \to F$. This is non-zero by construction. Hence the kernel $v_1^{\perp}$ is $(n-1)$-dimensional. By induction, $v_1^{\perp}$ has an orthogonal basis. Since $v_1 \notin v_1^{\perp}$, tacking on v_1 gives an orthogonal basis for V.

2) We will show $V = U \oplus U^{\perp}$, so any orthogonal basis of U can be extended to an orthogonal basis of V by using an orthogonal basis of $U^{\perp}$ (note $U \perp U^{\perp}$ by definition). Notice that $U \cap U^{\perp} = \{0\}$ because U is non-degenerate. Take any $v \in V$. Then $\mathrm{B}(v, -)|_U \in U^*$, so by Theorem 2.8 (applied to the vector space U) there is some $u_0 \in U$ such that $\mathrm{B}(v, u) = \mathrm{B}(u_0, u)$ for all $u \in U$. Then $\mathrm{B}(v - u_0, u) = 0$ for all $u \in U$. So $v - u_0 \in U^{\perp}$ and

$$v = u_0 + (v - u_0).$$

Thus, $V = U \oplus U^{\perp}$ as claimed. □

Corollary 2.23. *Let* B *be non-degenerate on* V. *If* $U \subset V$ *is a non-degenerate subspace, then* $U^{\perp}$ *is non-degenerate.*

Proof. Suppose there is some $v \in U^{\perp}$ such that for all $w \in U^{\perp}$, $\mathrm{B}(v, w) = 0$. By definition $v \perp U$ so since $V = U \oplus U^{\perp}$ by Theorem 2.22 we have $v \perp V$, forcing $v = 0$. □

We are now ready to return to the discussion of quadratic forms using symmetric bilinear forms to develop the theory in a way that is coordinate-free.

2.2 Quadratic Forms

In this section, we make the convention that F is *not* of characteristic 2.

Definition 2.24. A *quadratic form* on V is a function $Q : V \to F$ such that

1) $Q(cv) = c^2 Q(v)$ for any $c \in F$ and $v \in V$,

2) the pairing $\mathrm{B} : V \times V \to F$ given by

$$\mathrm{B}(v, w) = \frac{1}{2}(Q(v + w) - Q(v) - Q(w))$$

for $v, w \in V$, is bilinear.

Just as in the case of the inner product on $\mathbb{R}^n$, we can recover Q from B by looking at the diagonal:

$$\begin{aligned} \mathrm{B}(v, v) &= \frac{1}{2}(Q(2v) - 2Q(v)) \\ &= \frac{1}{2}(4Q(v) - 2Q(v)) \\ &= Q(v). \end{aligned}$$

It will often be convenient to use property (2) of Definition 2.24 in the form

$$Q(v + w) = Q(v) + Q(w) + 2\mathrm{B}(v, w).$$

More generally,

$$Q(v_1 + \cdots + v_r) = Q(v_1) + \cdots + Q(v_r) + 2\sum_{i<j} \mathrm{B}(v_i, v_j). \tag{2.2}$$

It is clear that B, as defined in Definition 2.24, is symmetric. Conversely, if B is any symmetric bilinear form on V then the function $Q(v) = \mathrm{B}(v, v)$ is a quadratic form whose associated bilinear form is B:

$$\begin{aligned} \frac{1}{2}(\mathrm{B}(v + w, v + w) - \mathrm{B}(v, v) - \mathrm{B}(w, w)) &= \frac{1}{2}(\mathrm{B}(v, w) + \mathrm{B}(w, v)) \\ &= \mathrm{B}(v, w). \end{aligned}$$

This calculation is referred to as the polarization identity.

From this point of view, quadratic forms and symmetric bilinear forms on V are equivalent concepts. Moreover, they both have linear structure which is respected by the correspondence between them. We also see that if $\mathrm{B}(v, w) = [v] \cdot A[w]$ in a basis, then $Q(v) = [v] \cdot A[v]$ since $Q(v) = B(v, v)$.

Example 2.25. The quadratic form $Q(x,y) = x^2 - y^2$ on $\mathbb{R}^2$ has the associated bilinear form $\mathrm{B}((x,y),(x',y')) = xx' - yy'$. In terms of matrices,

$$Q(x,y) = x^2 - y^2 = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and

$$\mathrm{B}((x,y),(x',y')) = xx' - yy' = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Let Q be a quadratic form on V and $v_1, \ldots, v_n$ a basis of V. Using (2.2) we can express Q as a homogeneous quadratic polynomial in these basis coordinates: for $x_1, \ldots, x_n \in F$,

$$\begin{aligned} Q(x_1v_1 + \cdots + x_nv_n) &= \sum_i Q(x_iv_i) + \sum_{i<j} 2\mathrm{B}(x_iv_i, x_jv_j) \\ &= \sum_i a_i x_i^2 + \sum_{i<j} a_{ij} x_i x_j \end{aligned}$$

where $a_i = Q(v_i)$ and $a_{ij} = 2\mathrm{B}(v_i, v_j)$. Write a_i as a_{ii} so

$$Q(x_1v_1 + \cdots + x_nv_n) = \sum_{i \le j} a_{ij} x_i x_j.$$

Then the column vector isomorphism $[\ ] : V \to F^n$ associated to $\{v_1, \ldots, v_n\}$ lets us write Q concretely in terms of a matrix:

$$Q(x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12}/2 & \cdots & a_{1n}/2 \\ a_{12}/2 & a_{22} & \cdots & a_{2n}/2 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}/2 & a_{2n}/2 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

where $\cdot$ is the standard dot product on F^n.

Definition 2.26. The *dimension* of a quadratic form Q on V, denoted $\dim Q$, is the dimension of V.

Definition 2.27. The *discriminant* of a quadratic form Q is the discriminant of its associated bilinear form and is denoted $\mathrm{disc}(Q)$.

In terms of a matrix representation $Q(v) = [v] \cdot A[v]$, the discriminant of Q is the determinant of A.

Definition 2.28. A quadratic form is *non-degenerate* when its associated symmetric bilinear form is non-degenerate in the sense of Definition 2.5.

Concretely, writing $Q(v) = [v] \cdot A[v]$ in some basis, Q is non-degenerate when A is invertible. So Q is non-degenerate if and only if $\mathrm{disc}(Q) \neq 0$.

Example 2.29. On $\mathbb{R}^2$, $Q(x, y) = x^2$ is degenerate: $\mathrm{B}((0, 1), v) = 0$ for all $v \in \mathbb{R}^2$. In terms of matrices,

$$Q(x, y) = \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is not invertible.

Remark 2.30. The non-zero values of a quadratic form $Q : V \to F$ are a set of square classes: for $c \in F^\times$ and $v \in V$, if $Q(v) \neq 0$ then $c^2 Q(v) = Q(cv)$. This means that we can think of the non-zero values of Q as lying in $F^\times / F^{\times^2}$. That is, if $Q(v)$ lies in a coset of $F^\times / F^{\times^2}$ then every other element of that coset is a value of Q. For instance, $1 \in Q(V)$ if and only if $F^{\times^2} \subset Q(V)$.

Definition 2.31. A non-zero vector $v \in V$ is called a *null vector* (or an *isotropic vector*) for Q if $Q(v) = 0$.

Notice that the zero vector is not a null vector (just as the zero vector is not considered to be an eigenvector of a linear map). A quadratic form has a null vector when the equation $Q(v) = 0$ is non-trivially solvable.

Theorem 2.32. *Any degenerate quadratic form has a null vector.*

Proof. If Q is degenerate then its associated symmetric bilinear form is degenerate. This means there is some non-zero $v \in V$ such that $\mathrm{B}(v, w) = 0$ for all $w \in V$. In particular, $\mathrm{B}(v, v) = Q(v) = 0$. □

Definition 2.33. For a quadratic form Q, the vectors $v, w \in V$ are called *orthogonal* if $\mathrm{B}(v, w) = 0$, where B is the bilinear form associated to Q.

Orthogonality of vectors gives a different perspective on null vectors: v is a null vector for Q if and only if $v \perp v$ and $v \neq 0$. It also can be described

directly in terms of Q since we always have $Q(v+w) = Q(v)+Q(w)+2\mathrm{B}(v,w)$: $v \perp w$ if and only if $Q(v+w) = Q(v)+Q(w)$.

Thus in an orthogonal basis $\{e_1, \ldots, e_n\}$ (which exists by Theorem 2.22),

$$\begin{aligned} Q(x_1e_1 + \cdots + x_ne_n) &= x_1^2Q(e_1) + \cdots + x_n^2Q(e_n) \\ &= a_1x_1^2 + \cdots + a_nx_n^2 \end{aligned} \tag{2.3}$$

where $a_i = Q(e_i)$. Any non-zero value of Q can occur as a coefficient of Q in some diagonalization. Indeed, start with any non-null vector e_1 and extend it to an orthogonal basis. Then (2.3) shows that $Q(e_1) = a_1$ occurs as the first coefficient of that diagonalization.

In an orthogonal basis, the matrix representation is diagonal so the discriminant is just the product of the diagonal terms.

Example 2.34. If $Q(x,y,z) = x^2 + y^2 + 3z^2$ then $\mathrm{disc}(Q) = 3$.

Furthermore, if Q is non-degenerate then a null vector cannot be part of an orthogonal basis since then it would be orthogonal to the whole space. Any non-null vector, however, is part of an orthogonal basis (regardless of whether or not Q is non-degenerate) by letting U in Theorem 2.22 be the span of the non-null vector.

Example 2.35. The trace form

$$Q(x,y,z,t) = 4x^2 - 36z^2 + 192t^2 + 48xt + 48yz - 72yt$$

from Example 1.13 can be diagonalized over $\mathbb{Q}$ in the following manner. We construct an orthogonal basis starting with the vector $(1,0,0,0)$. Notice that $Q(1,0,0,0) = 4$ so it is not a null vector. Next use the matrix realization of the bilinear form B attached to Q to find constraints on other vectors that could make up the orthogonal basis. In this case, the matrix realization gives

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 0 & 24 \\ 0 & 0 & 24 & -36 \\ 0 & 24 & -36 & 0 \\ 24 & -36 & 0 & 192 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 4x + 24t. \tag{2.4}$$

For a vector v to be included in an orthogonal basis with $(1,0,0,0)$, we need $Q(v) \neq 0$ while v makes the right side of (2.4) vanish. For example,

$(0,1,0,0)$ does not work because $Q(0,1,0,0) = 0$, but $(0,0,1,0)$ does work: (2.4) vanishes at (0,0,1,0) and $Q(0,0,1,0) = -36$. Then

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 0 & 24 \\ 0 & 0 & 24 & -36 \\ 0 & 24 & -36 & 0 \\ 24 & -36 & 0 & 192 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = 24y - 36z \qquad (2.5)$$

and a non-zero vector that makes (2.4) and (2.5) vanish is $(0,3,2,0)$. Furthermore, $Q(0,3,2,0) = 144$. The condition for a vector to be orthogonal to $(0,3,2,0)$ is $48y - 108t = 0$. This equation is satisfied by $(-24,9,6,4)$ which also makes (2.4) and (2.5) vanish. Also $Q(-24,9,6,4) = -528 = -16 \cdot 33$, so

$$\{(1,0,0,0),(0,0,1,0),(0,3,2,0),(-24,9,6,4)\}$$

is an orthogonal basis for Q. In this basis Q is diagonalized:

$$Q(x,y,z,t) = 4x^2 - 36y^2 + 144z^2 - 528t^2.$$

Definition 2.36. A quadratic form Q on V is called *universal* if $Q(V) = F$.

Example 2.37. The quadratic form $x^2 - y^2$ on F^2 is universal over F: for any $a \in F, a = (\frac{a+1}{2})^2 - (\frac{a-1}{2})^2$.

Theorem 2.38. *Let $a \in F^\times$. The quadratic form $Q(x,y) = x^2 - ay^2$ on F^2 has a null vector if and only if a is a square in $F^\times$, in which case Q is universal.*

Proof. If a is a square in $F^\times$, say $a = c^2$, then $(c,1)$ is a null vector for Q. Conversely, say $(x_0,y_0) \in F^2$ is a null vector for Q. So $x_0^2 = ay_0^2$ and $(x_0,y_0) \neq (0,0)$. Then in fact x_0 and y_0 are both non-zero (since $a \neq 0$) so $a = (x_0/y_0)^2$ is a square in $F^\times$. Thus $Q(x,y) = x^2 - (cy)^2$ where $c = x_0/y_0$ and this is universal by the calculation in Example 2.37. □

Theorem 2.39. *Let Q be a non-degenerate quadratic form on V. If Q has a null vector then Q is universal.*

Proof. We offer two proofs. For the first proof, let $v \in V$ be a null vector for Q. Since the bilinear form B associated to Q is non-degenerate, there is some $w \in V$ such that $\mathrm{B}(v,w) \neq 0$. Then for any $c \in F$,

$$\begin{aligned} Q(cv+w) &= Q(cv) + Q(w) + 2B(cv,w) \\ &= c^2Q(v) + Q(w) + 2cB(v,w) \\ &= Q(w) + 2cB(v,w). \end{aligned} \qquad (2.6)$$

Since $2\mathrm{B}(v,w)$ is non-zero (we are not in characteristic 2), (2.6) takes on all values in F as c runs over F. So Q is universal. This concludes the first proof.

For our second proof, we can assume $\dim V \geq 2$ since a one-dimensional non-degenerate quadratic form cannot have a null vector.

Note that for any vectors $v, w \in V$ and scalars $\alpha, \beta \in F$,

$$\begin{aligned} Q(\alpha v + \beta w) &= Q(\alpha v) + Q(\beta w) + 2\mathrm{B}(\alpha v, \beta w) \\ &= \alpha^2 Q(v) + \beta^2 Q(w) + 2\alpha\beta \mathrm{B}(v,w). \end{aligned} \tag{2.7}$$

Let v_0 be a null vector for Q. If we can find another null vector w_0 such that $\mathrm{B}(v_0, w_0) \neq 0$, then v_0 and w_0 are linearly independent (if $w_0 = cv_0$, then $\mathrm{B}(v_0, w_0) = c\mathrm{B}(v_0, v_0) = 0$). Moreover, we would then have

$$Q(\alpha v_0 + \beta w_0) = 2\alpha\beta \mathrm{B}(v_0, w_0). \tag{2.8}$$

Thus, if we fix $\beta = 1$ and let α vary in F (just like Equation (2.6) in the first proof) then the right side of (2.8) takes on all values in F, so Q is universal.

We will now find such a w_0. By non-degeneracy, $\mathrm{B}(v_0, v_0') \neq 0$ for some v_0'. Then by scaling v_0' we can assume $\mathrm{B}(v_0, v_0') = 1$. So $\{v_0, v_0'\}$ is linearly independent and

$$Q(\alpha v_0 + v_0') = Q(v_0') + 2\alpha \tag{2.9}$$

by (2.7). The right side of (2.9) is zero when $\alpha = -Q(v_0')/2$ (remember that $\mathrm{char}(F) \neq 2$). Therefore we can use $w_0 = -(Q(v_0')/2)v_0 + v_0'$ as our second null vector. □

Remark 2.40. The non-degeneracy condition is crucial in Theorem 2.39. For example, the quadratic form on $\mathbb{R}^2$ in Example 2.29 has $(0,1)$ as a null vector but it is not universal on $\mathbb{R}^2$.

Remark 2.41. In the second proof of Theorem 2.39, we can scale w_0 so $\mathrm{B}(v_0, w_0) = 1$. Then $Q(v_0 + \frac{1}{2}w_0) = 1$ and $Q(v_0 - \frac{1}{2}w_0) = -1$. So on the plane spanned by $v_0 + \frac{1}{2}w_0, v_0 - \frac{1}{2}w_0$,

$$Q(x(v_0 + \frac{1}{2}w_0) + y(v_0 - \frac{1}{2}w_0)) = x^2 - y^2.$$

This means that Q has a null vector if and only if V contains a plane on which Q is $x^2 - y^2$.

What about the converse of Theorem 2.39? If a quadratic form is non-degenerate and universal, does this always imply that it has a null vector? No!

Example 2.42. Let $F = \mathbb{F}_p$ with $p \neq 2$. Let a be a non-square in $\mathbb{F}_p$. Then the quadratic form

$$\begin{aligned} Q(x,y) &= x^2 - ay^2 \\ &= \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

is non-degenerate and it does not have a null vector by Theorem 2.38. To see that it is universal, notice that it is the norm map N from $\mathbb{F}_{p^2} = \mathbb{F}_p(\sqrt{a})$ to $\mathbb{F}_p$ and $\mathrm{N} : \mathbb{F}_{p^2}^\times \to \mathbb{F}_p^\times$ is onto (a generator goes to a generator) while $0 = \mathrm{N}(0)$.

Corollary 2.43. *Let $Q(x_1, \ldots, x_n)$ be a non-degenerate quadratic form on F^n. For $r \in F^\times$, the following are equivalent:*

1) *$Q(a_1, \ldots, a_n) = r$ is solvable for some $(a_1, \ldots, a_n) \in F^n$,*

2) *the quadratic form $Q(x_1, \ldots, x_n) - rx_{n+1}^2$ on F^{n+1} has a null vector.*

Proof. Suppose (1) is true: $Q(a_1, \ldots, a_n) = r$. Then

$$Q(a_1, \ldots, a_n) - rx_{n+1}^2 = r - rx_{n+1}^2,$$

so $(a_1, \ldots, a_n, 1)$ is a null vector for $Q(x_1, \ldots, x_n) - rx_{n+1}^2$.

Conversely, suppose (2) is true. Let $(a_1, \ldots, a_n, a_{n+1})$ be a null vector. If $a_{n+1} \neq 0$, then

$$r = \frac{1}{a_{n+1}^2} Q(a_1, \ldots, a_n) = Q\left(\frac{a_1}{a_{n+1}}, \ldots, \frac{a_n}{a_{n+1}}\right).$$

So Q represents r in F^n.

If $a_{n+1} = 0$, then $(a_1, \ldots, a_n) \neq \mathbf{0}$ since $(a_1, \ldots, a_n, a_{n+1}) \neq \mathbf{0}$. Then Q has a null vector $(a_1, \ldots, a_n)$, so Q is universal on F^n by Theorem 2.39. In particular, Q represents r in F^n. □

Example 2.44. We can now describe a counterexample to the converse of Theorem 2.39 over $\mathbb{Q}$. The quadratic form $x^2 + y^2 + z^2 - 7t^2$ on $\mathbb{Q}^4$ does not have a null vector: if it did, then 7 is a sum of three rational squares by Corollary 2.43. So 7 is a sum of three integer squares by Theorem 1.6, but this is false. We will see in Section 4.5, however, that the quadratic form is universal on $\mathbb{Q}^4$ by the Hasse–Minkowski theorem.

Corollary 2.45. *Let Q_1 and Q_2 be non-degenerate quadratic forms on V. Set Q to be the following quadratic form on $V \oplus V$:*

$$Q(v, w) = Q_1(v) - Q_2(w).$$

Then Q_1 and Q_2 have a common non-zero value on V if and only if Q has a null vector on V.

Proof. If Q_1 and Q_2 have a common non-zero value, then it is clear that Q has a null vector.

Conversely, first suppose that Q_1 or Q_2 has a null vector v_0. Then $(v_0, 0)$ or $(0, v_0)$ is a null vector of Q. This means that at least one of Q_1, Q_2 is universal since they are non-degenerate, so without loss of generality let Q_1 be universal. Then Q_1 and Q_2 represent a common non-zero value over F since Q_2 is non-degenerate and, therefore, takes on *some* non-zero values over F. Notice that Q having a null vector plays a trivial role in this case.

Now suppose that Q has a null vector and neither Q_1 nor Q_2 have null vectors. Since Q has a null vector, say (v_0, w_0), we have

$$Q_1(v_0) = Q_2(w_0) \tag{2.10}$$

and $(v_0, w_0) \neq (0, 0)$. The common value of Q_1 and Q_2 in (2.10) is non-zero since $Q_1(v_0) \neq 0$ if $v_0 \neq 0$ and $Q_2(w_0) \neq 0$ if $w_0 \neq 0$. □

Definition 2.46. A *quadratic space* (over F) is a pair (V, Q) where Q is a quadratic form on V.

Example 2.47. Let K/F be a quadratic field extension. Then viewing K as an F-vector space, $(K, \mathrm{N}_{K/F})$ is a non-degenerate quadratic space: writing $K = F(\sqrt{a})$, we have $\mathrm{N}_{K/F}(x + y\sqrt{a}) = x^2 - ay^2$.

Definition 2.48. Let (V, Q) be a quadratic space. Two subspaces $W_1, W_2 \subset V$ are called *orthogonal*, and we write $W_1 \perp W_2$, if $\mathrm{B}(w_1, w_2) = 0$ for all $w_1 \in W_1$ and $w_2 \in W_2$, where B is the symmetric bilinear form associated to Q. A subspace $W \subset V$ is *non-degenerate* if it is non-degenerate for this B.

Remark 2.49. For $v \in V$, $Q(v) \neq 0$ if and only if the line Fv is a non-degenerate subspace of V.

Definition 2.50. Let (V, Q) be a quadratic space. If $V = W_1 \oplus \cdots \oplus W_r$ and $W_i \perp W_j = 0$ for $i \neq j$, then we say V is an *orthogonal direct sum* of $W_1, \ldots, W_r$. Denote this by $V = W_1 \perp \cdots \perp W_r$.

Although there is a slight abuse of the notation $\perp$ for subspaces since it is both a *relation* (orthogonal subspaces) and a *construction* (orthogonal direct sum of subspaces), the context in which it is used should make the meaning evident.

Example 2.51. If $Q(x, y) = x^2 + xy - y^2$ on $\mathbb{R}^2$, then the usual direct sum decomposition $\mathbb{R}^2 = \mathbb{R}(1, 0) \oplus \mathbb{R}(0, 1)$ is not an orthogonal direct sum with respect to Q because $(1, 0) \not\perp (0, 1)$. An orthogonal direct sum decomposition of $(\mathbb{R}^2, Q)$ is $\mathbb{R}(1, 0) \oplus \mathbb{R}(1, -2)$.

Example 2.52. Any quadratic space (V, Q) is an orthogonal direct sum of lines:

$$V = L_1 \perp \cdots \perp L_d$$

since any V has an orthogonal basis by Theorem 2.22 (each L_i is the span of a basis vector).

Just as we defined orthogonal direct sums of subspaces, we can also consider building a quadratic space out of quadratic spaces of smaller dimension. That is, if $(V_1, Q_1), \ldots, (V_r, Q_r)$ are quadratic spaces, then $W = V_1 \oplus \cdots \oplus V_r$ is also a quadratic space with the quadratic form $Q(v_1, \ldots, v_r) = \sum_{i=1}^{r} Q_i(v_i)$. It is clear that V_i and V_j are orthogonal subspaces in W for $i \neq j$ since there is no interaction between the quadratic forms Q_i and Q_j for $i \neq j$. We write $W = V_1 \perp \cdots \perp V_r$.

Definition 2.53. Quadratic forms Q_1 on V_1 and Q_2 on V_2 are *equivalent* when their associated bilinear forms are equivalent. We then write $Q_1 \cong Q_2$.

Theorem 2.54. *For quadratic forms Q_1 on V_1 and Q_2 on V_2, $Q_1 \cong Q_2$ if and only if there is a linear isomorphism $C : V_1 \to V_2$ such that $Q_2(Cv) = Q_1(v)$ for all $v \in V_1$.*

Proof. Suppose $Q_1 \cong Q_2$. By definition, this means that there is a linear isomorphism $C : V_1 \to V_2$ such that $\mathrm{B}_2(Cv, Cw) = \mathrm{B}_1(v, w)$ for all $v, w \in V_1$. Then

$$Q_2(Cv) = \mathrm{B}_2(Cv, Cv) = \mathrm{B}_1(v, v) = Q_1(v).$$

for all $v \in V_1$.

Conversely, suppose there is a linear isomorphism $C : V_1 \to V_2$ such that $Q_2(Cv) = Q_1(v)$ for all $v \in V_1$. Then by polarization $\mathrm{B}_2(Cv, Cw) = \mathrm{B}_1(v, w)$ for all $v, w \in V_1$. Hence $Q_1 \cong Q_2$. $\square$

In other words, the quadratic forms Q_1 and Q_2 in Theorem 2.54 are equivalent if there are bases in V_1 and V_2 where Q_1 and Q_2 are the same homogeneous quadratic polynomial.

Example 2.55. The trace form

$$4x^2 - 36z^2 + 192t^2 + 48xt + 48yz - 72yt$$

from Example 1.13 is equivalent over $\mathbb{Q}$ to the quadratic form

$$4x^2 - 36y^2 + 144z^2 - 528t^2 \tag{2.11}$$

from Example 2.35. After scaling the variables, (2.11) is equivalent to

$$x^2 - y^2 + z^2 - 33t^2$$

over $\mathbb{Q}$. This last quadratic form trivially has a null vector over $\mathbb{Q}$ so it is universal by Theorem 2.39. In particular, it takes on the value 3 using $x = 2, y = 1$ and $z = t = 0$. Since it is equivalent over $\mathbb{Q}$ to the trace form in Example 1.13 we now know that this trace form not only takes on the value 3 over $\mathbb{Q}$, but it is also universal over $\mathbb{Q}$.

Definition 2.56. Let (V_1, Q_1) and (V_2, Q_2) be quadratic spaces. A *quadratic space isomorphism* $\varphi : V_1 \to V_2$ is a linear isomorphism of vector spaces that respects the quadratic form structure on the vector spaces. That is, $Q_2(\varphi(v)) = Q_1(v)$ for all $v \in V_1$.

Essentially, saying quadratic spaces are isomorphic means the same thing as saying the underlying quadratic forms are equivalent. In other words, it is just a matter of terminology.

Now we look at the 2-dimensional case. Recall that when $x = yz^2$ for $x, y \in F$ and $z \in F^\times$, we denote this by $x \sim y$.

Lemma 2.57. *Let (V_1, Q_1) and (V_2, Q_2) be non-degenerate 2-dimensional quadratic spaces. Then $Q_1 \cong Q_2$ if and only if $\operatorname{disc}(Q_1) \sim \operatorname{disc}(Q_2)$ and there is some $c \in F^\times$ such that $c \in Q_1(V_1) \cap Q_2(V_2)$.*

Proof. It is clear that if $Q_1 \cong Q_2$ then $\operatorname{disc}(Q_1) \sim \operatorname{disc}(Q_2)$ and by Theorem 2.54, $Q_1(V_1) = Q_2(V_2)$. In particular, there is a non-zero $c \in Q_1(V_1) \cap Q_2(V_2)$. Notice that the constraint on the dimension is not needed here: equivalent

quadratic forms have the same discriminant (up to non-zero squares) and the same set of values.

Conversely, suppose that $\text{disc}(Q_1) \sim \text{disc}(Q_2)$ and that there is some non-zero $c \in Q_1(V_1) \cap Q_2(V_2)$. Let $c = Q_1(v_1)$ for $v_1 \in V_1$. So $v_1 \neq 0$. From Theorem 2.22 (2), v_1 is part of an orthogonal basis $\{v_1, w_1\}$ of V_1. In this basis,

$$Q_1(xv_1 + yw_1) = cx^2 + dy^2,$$

where $d = Q_1(w_1)$. The same argument applies to Q_2. Writing $c = Q_2(v_2)$, there is an orthogonal basis $\{v_2, w_2\}$, so

$$Q_2(xv_2 + yw_2) = cx^2 + d'y^2,$$

where $d' = Q_2(w_2)$. Note that $d, d' \in F^\times$ since Q_1 and Q_2 are non-degenerate. Then computing the discriminants gives $cd = cd'a^2$ for some $a \in F^\times$ since $\text{disc}(Q_1) \sim \text{disc}(Q_2)$. Hence

$$\begin{aligned} Q_1(xv_1 + yw_1) &= cx^2 + dy^2 \\ &= cx^2 + d'a^2y^2 \\ &= Q_2(xv_2 + yaw_2), \end{aligned}$$

so $Q_1 \cong Q_2$ using the linear isomorphism $V_1 \to V_2$ given by $v_1 \mapsto v_2$ and $w_1 \mapsto aw_2$. □

Example 2.58. In Lemma 2.57 the condition that Q_1 and Q_2 take on a common value really is necessary. If we dropped this condition then the lemma would imply that $x^2 + y^2$ is equivalent to $3x^2 + 3y^2$ over $\mathbb{Q}$ which means they take on the same set of values over $\mathbb{Q}$. So $x^2 + y^2 = 3$ for some $x, y \in \mathbb{Q}$ and thus $x^2 + y^2 = 3$ for some $x, y \in \mathbb{Z}$ by Theorem 1.5, but this is false.

Definition 2.59. A *hyperbolic plane* is a two-dimensional quadratic space (V, Q) where, in some basis $\{e_1, e_2\}$, $Q(xe_1 + ye_2) = x^2 - y^2$.

All hyperbolic planes are equivalent. We denote a hyperbolic plane as H. Any hyperbolic plane contains a null vector (for instance, $e_1 + e_2$ using the notation from Definition 2.59). Conversely, from Remark 2.41, any non-degenerate quadratic space (V, Q) that contains a null vector has $\dim V \geq 2$ and also contains a hyperbolic plane. In particular, when $\dim V = 2$, (V, Q) is a hyperbolic plane if and only if it is non-degenerate and contains a null vector.

Theorem 2.60. *For a two-dimensional quadratic space (V, Q), the following are equivalent:*

1) *(V, Q) is a hyperbolic plane,*

2) *(V, Q) is non-degenerate and contains a null vector,*

3) $\text{disc}(Q) \sim -1$.

Proof. That (2) $\Rightarrow$ (1) comes from Remark 2.41. Now let (V, Q) be a hyperbolic plane. By definition, $Q(xe_1 + ye_2) = x^2 - y^2$ in some basis $\{e_1, e_2\}$. So $\text{disc}(Q) \sim -1$. Thus (1) $\Rightarrow$ (3).

Lastly, we will show (3) $\Rightarrow$ (2). The quadratic space (V, Q) is non-degenerate since $\text{disc}(Q) \neq 0$. Choose $v \in V$ such that $Q(v) \neq 0$. Extend v to an orthogonal basis $\{v, w\}$ of V. Relative to the basis $\{v, w\}$,

$$Q(xv + yw) = Q(v)x^2 + Q(w)y^2 = cx^2 + Q(w)y^2,$$

where $c = Q(v)$. Then $\text{disc}(Q) = cQ(w) \sim -1$ by our hypothesis, which means

$$Q(w) \sim \frac{-1}{c} \sim -c$$

since $c \neq 0$. Then by scaling w, we can assume $Q(w) = -c$. Thus, $Q(v+w) = c - c = 0$. So $v + w$ is a null vector. □

Remark 2.61. In the proof of Theorem 2.60, that $v + w$ is a null vector corresponds to the obvious null vector $(1, 1)$ for $x^2 - y^2$ on F^2.

Theorem 2.62. *If (V, Q) is non-degenerate and contains a null vector then $V \cong H \perp W$ where H is a hyperbolic plane and W is non-degenerate.*

Proof. From the proof of Theorem 2.39 and from Remark 2.41 we know that V contains a hyperbolic plane H. By definition H is non-degenerate so $V = H \perp H^{\perp}$ and $H^{\perp}$ is non-degenerate (Corollary 2.23). □

Theorem 2.63. *Let (V, Q) be a four-dimensional non-degenerate quadratic space. If Q has a null vector in V and $\text{disc}(Q) \sim 1$, then any non-degenerate three-dimensional subspace has a null vector.*

Proof. Let $U \subset V$ be a non-degenerate three-dimensional subspace of V. Then $U^\perp$ is one-dimensional since V is non-degenerate and $V = U \perp U^\perp$ by Theorem 2.22. Let $v \in U^\perp$ be non-zero. Then $V = Fv \perp U$ and $Q(v) \neq 0$ since $U^\perp$ is non-degenerate. We claim that $Q(v) = -Q(u)$ for some $u \in U$.

Suppose U does not have a null vector since otherwise we are done (Q would be universal on U). This means that V has a null vector $cv + u'$ for some $c \in F^\times$ and $u' \in U$. So

$$0 = Q(cv + u') = c^2 Q(v) + Q(u'). \tag{2.12}$$

Thus (2.12) implies $Q(v) = -Q(u'/c)$.

Let $u = u'/c$. Since $Q(u) \neq 0$, there is an orthogonal basis of U containing u (Theorem 2.22). Hence $V = Fv \perp Fu \perp W$ where $W \subset U$ is a non-degenerate two-dimensional subspace of U. So

$$1 \sim \operatorname{disc} V = -Q(v)^2 \operatorname{disc} W \sim -\operatorname{disc} W.$$

Then $\operatorname{disc} W \sim -1$ so W has a null vector by Theorem 2.60, which means U has a null vector. □

Remark 2.64. In Theorem 2.63, $V \cong H \perp V'$ where H is a hyperbolic plane and V' is non-degenerate by Theorem 2.62. So

$$\operatorname{disc}(H) \operatorname{disc}(V') \sim 1$$

since $\operatorname{disc}(V) \sim 1$. This means $\operatorname{disc}(V') \sim -1$ and thus $V' \cong H$ by Theorem 2.60. So $V \cong H \perp H$.

Remark 2.65. Theorem 2.63 does not extend to two-dimensional subspaces. For example, the quadratic form $Q(x, y, z, t) = x^2 + y^2 - z^2 - t^2$ on $\mathbb{R}^4$ is non-degenerate, its discriminant is 1 and it has $(1, 0, 1, 0)$ as a null vector. The non-degenerate subspace spanned by the vectors $e_1 = (1, 0, 0, 0)$ and $e_2 = (0, 1, 0, 0)$, however, does not have a null vector: $Q(xe_1 + ye_2) = x^2 + y^2$.

In Example 2.47 we noted that the norm map for a quadratic field extension K/F is a non-degenerate quadratic form. The next result is an analogue of Theorem 2.60 when there is no null vector.

Theorem 2.66. *For a 2-dimensional quadratic space (V, Q) over F, the following are equivalent:*

1) $(V, Q) \cong (K, \mathrm{N}_{K/F})$ *where* K/F *is a quadratic field extension,*

2) (V, Q) *is non-degenerate, does not have a null vector, and* $1 \in Q(V)$*,*

3) $\mathrm{disc}(Q) \not\sim 0$*,* $\mathrm{disc}(Q) \not\sim -1$ *and* $1 \in Q(V)$*.*

Proof. (1) $\Rightarrow$ (2): This comes from Theorem 2.38.

(2) $\Rightarrow$ (3): Since Q is non-degenerate, $\mathrm{disc}(Q) \not\sim 0$. Also $\mathrm{disc}(Q) \not\sim -1$ because otherwise it would have a null vector (Theorem 2.60).

(3) $\Rightarrow$ (1): Since $\mathrm{disc}(Q) \not\sim 0$, $\mathrm{disc}(Q) \not\sim -1$ and $1 \in Q(V)$ we have that $Q \cong x^2 + ay^2$ for a non-square $-a$ in $F^\times$. Hence $Q \cong \mathrm{N}_{K/F}$ where $K = F(\sqrt{-a})$. □

This section concludes with a discussion of some transformations that preserve the "geometry" of quadratic forms and an important application of them to quadratic spaces (Theorem 2.69).

Definition 2.67. If $v \in V$ and $Q(v) \neq 0$, the *reflection* $\tau_v : V \to V$ is the function defined by

$$\tau_v(w) = w - 2\frac{\mathrm{B}(v, w)}{Q(v)}v.$$

Here are some properties of reflections which come directly out of Definition 2.67.

- A reflection τ_v is linear since $w \mapsto w$ and $w \mapsto \mathrm{B}(v, w)$ are linear.
- On $v^\perp$, τ_v is the identity function since $\mathrm{B}(v, v') = 0$ for $v' \in v^\perp$.
- Applying τ_v to any $cv \in Fv$, we see

  $$\tau_v(cv) = cv - 2\frac{\mathrm{B}(v, cv)}{Q(v)}v = -cv.$$

  So cv is "reflected" across the hyperplane $v^\perp$. Notice that this also means τ_v^2 is the identity map on all of V since $V = Fv \perp v^\perp$ (recall that Fv is non-degenerate because $Q(v) \neq 0$).
- The reflection τ_v is a linear automorphism since τ_v is its own inverse.

Conversely, for any $v \in V$ where $Q(v) \neq 0$ assume there is a linear automorphism τ of V such that $\tau(v) = -v$ and $\tau(v') = v'$ for $v' \in v^\perp$. Then $V = Fv \perp v^\perp$ by Theorem 2.22 so any w in V has the form $w = av + v'$ for some $a \in F$ and $v' \in v^\perp$. Then $\mathrm{B}(v, w) = \mathrm{B}(v, av) = aQ(v)$, so

$$\begin{aligned}
\tau(w) &= \tau(av + v') \\
&= a\tau(v) + \tau(v') \\
&= -av + v' \\
&= av + v' - 2av \\
&= av + v' - 2\frac{\mathrm{B}(v, w)}{Q(v)}v \\
&= w - 2\frac{\mathrm{B}(v, w)}{Q(v)}v,
\end{aligned}$$

so $\tau = \tau_v$. We have derived the definition of the reflection τ_v from the properties that it has.

Another property of reflections is that τ_v is an *isometry* in the sense that $Q(\tau_v(u)) = Q(u)$ for all $u \in V$ and, more generally, $\mathrm{B}(\tau_v(u), \tau_v(w)) = \mathrm{B}(u, w)$ for all $u, w \in V$. We check these by a direct calculation:

$$\begin{aligned}
Q(\tau_v(u)) &= \mathrm{B}(\tau_v(u), \tau_v(u)) \\
&= \mathrm{B}\left(u - 2\frac{\mathrm{B}(v, u)}{Q(v)}v, u - 2\frac{\mathrm{B}(v, u)}{Q(v)}v\right) \\
&= \mathrm{B}(u, u) - 4\frac{\mathrm{B}(v, u)^2}{Q(v)} + 4\frac{\mathrm{B}(v, u)^2}{Q(v)} \\
&= Q(u).
\end{aligned}$$

To see that $\mathrm{B}(\tau_v(u), \tau_v(w)) = \mathrm{B}(u, w)$ use polarization. That is,

$$\begin{aligned}
\mathrm{B}(\tau_v(u), \tau_v(w)) &= \frac{1}{2}(Q(\tau_v(u) + \tau_v(w)) - Q(\tau_v(u)) - Q(\tau_v(w))) \\
&= \frac{1}{2}(Q(\tau_v(u + w)) - Q(\tau_v(u)) - Q(\tau_v(w))) \\
&= \frac{1}{2}(Q(u + w) - Q(u) - Q(w)) \\
&= \mathrm{B}(u, w).
\end{aligned}$$

These reflections τ_v associated to non-null vectors of Q are analogous to reflections in $\mathbb{R}^n$ across $(n-1)$-dimensional hyperplanes. Notice that a reflection on a quadratic space (V, Q) is an isomorphism of (V, Q) with itself.

Any reflection τ is an isometry ($Q(\tau(w)) = Q(w)$ for all $w \in V$)) so it is natural to ask if $v = \tau(w)$ for some reflection τ when $Q(v) = Q(w)$. As it turns out, one reflection is not always sufficient for carrying w to v. However, when v and w are not null vectors, w can be mapped to v using at most two reflections.

Theorem 2.68. *Let v and w be vectors in V such that $Q(v) = Q(w) \neq 0$. Then $v = \tau(w)$ or $v = \tau_1\tau_2(w)$ for some reflections τ, τ_1, τ_2.*

Proof. If a reflection τ_x is going to map w to v, then

$$w - 2\frac{\mathrm{B}(x,w)}{Q(x)}x = v.$$

So if

$$2\frac{\mathrm{B}(x,w)}{Q(x)} = 1,$$

then $x = w - v$. On the other hand, if $Q(w-v) \neq 0$ then

$$\begin{aligned} 2\frac{\mathrm{B}(w-v,w)}{Q(w-v)} &= \frac{2Q(w) - 2\mathrm{B}(v,w)}{Q(w)+Q(v)-2\mathrm{B}(v,w)} \\ &= 1 \end{aligned}$$

since $Q(v) = Q(w)$. Thus, $\tau_{w-v}(w) = v$ if $Q(w-v) \neq 0$. We now check to see if $Q(w-v) \neq 0$.

Consider the sum

$$\begin{aligned} Q(w-v) + Q(w+v) &= 2Q(w) + 2Q(v) - 2\mathrm{B}(v,w) + 2\mathrm{B}(v,w) \\ &= 4Q(w) \neq 0. \end{aligned}$$

So one of $Q(w-v)$ and $Q(v+w)$ must be non-zero. If $Q(w-v) \neq 0$, then we are done. Otherwise, $Q(w+v) \neq 0$. In this case,

$$\begin{aligned} \tau_{w+v}(w) &= w - 2\frac{\mathrm{B}(w+v,w)}{Q(w+v)}(w+v) \\ &= w - \frac{2Q(w)+2\mathrm{B}(v,w)}{Q(w)+Q(v)+2\mathrm{B}(v,w)}(w+v) \\ &= w - (w+v) \\ &= -v. \end{aligned}$$

Therefore the composite of τ_v with τ_{w+v} sends w to $(\tau_v \circ \tau_{w+v})(w) = v$. So at most two reflections are needed to take w to v. □

The following theorem uses reflections to show that a common quadratic space in isomorphic orthogonal direct sums can be cancelled.

Theorem 2.69 (Witt cancellation)**.** *Let (V_i, Q_i) for $i = 1, 2, 3$ be three non-degenerate quadratic spaces. If the quadratic spaces $V_1 \perp V_2$ and $V_1 \perp V_3$ are isomorphic, then $V_2 \cong V_3$.*

Proof. Set $Q = Q_1 + Q_2$ and $Q' = Q_1 + Q_3$. There is $v \in V_1$ such that $Q(v) \neq 0$. Then Fv is a non-degenerate subspace so $V_1 = Fv \perp v^\perp$ where $\dim v^\perp = \dim V_1 - 1$. Let $V_1' = v^\perp$, so V_1' is non-degenerate. Then since the orthogonal sum is associative,

$$V_1 \perp V_2 = Fv \perp (V_1' \perp V_2)$$

and $V_1 \perp V_3 = Fv \perp (V_1' \perp V_3)$. So if the one-dimensional case of our theorem is true generally, then $V_1' \perp V_2 \cong V_1' \perp V_3$. Thus, since V_1 is an orthogonal direct sum of non-degenerate lines, successively cancelling non-degenerate one-dimensional subspaces gives the desired result.

So we may assume $\dim V_1 = 1$. Let $\varphi : (V_1 \perp V_2, Q) \to (V_1 \perp V_3, Q')$ be a quadratic space isomorphism: φ is linear and $Q'(\varphi v) = Q(v)$ for all $v \in V_1 \perp V_2$. Let B and B′ be the bilinear forms associated to Q and Q' respectively. Then using the relation $Q(v) = \mathrm{B}(v, v)$ and polarization we get $\mathrm{B}'(\varphi v, \varphi w) = \mathrm{B}(v, w)$ for all $v, w \in V_1 \perp V_2$.

Let $v_1 \in V_1$ be non-zero, so $V_1 = Fv_1$ and $Q_1(v_1) \neq 0$. Note $Q(v_1, v_2) = Q_1(v_1) + Q_2(v_2)$ and $Q'(v_1, v_3) = Q_1(v_1) + Q_3(v_3)$ for $v_i \in V_i$. Then by construction we have

$$Q'(\varphi v_1) = Q(v_1) = Q_1(v_1) = Q'(v_1) \neq 0.$$

So by Theorem 2.68 there is a reflection (or a composition of two reflections) τ on $V_1 \perp V_3$ such that $\tau(\varphi v_1) = v_1$. Thus, replacing φ with $\tau\varphi$, we may assume that $\varphi v_1 = v_1$. Then $\mathrm{B}'(v_1, \varphi w) = \mathrm{B}'(\varphi v_1, \varphi w) = \mathrm{B}(v_1, w)$ for all $w \in V_2$, so

$$\begin{aligned} \varphi w \in V_3 &\iff \varphi w \perp v_1 \text{ in } V_1 \perp V_3 \\ &\iff w \perp v_1 \text{ in } V_1 \perp V_2 \\ &\iff w \in V_2, \end{aligned}$$

which means $\varphi(V_2) \subset V_3$. Then $\varphi : V_2 \to V_3$ is an isomorphism of vector spaces because $\dim V_2 = \dim V_3$ and φ is injective. Restricting φ to V_2 gives an isomorphism from $(V_2, Q|_{V_2})$ to $(V_3, Q'|_{V_3})$, so $V_2 \cong V_3$. □

Example 2.70. The quadratic form

$$Q(x, y, z, t) = 4x^2 - 36y^2 + 144z^2 - 528t^2 \cong x^2 - y^2 + z^2 - 33t^2$$

from Example 2.35 is equivalent to

$$Q(x, y, z, t) = 222x^2 - 3y^2 + z^2 + 11t^2 + 36xy - 122xt - 20yt$$

on $\mathbb{Q}^4$ using the basis $\{(0, -1, 16, 1), (0, 2, 1, 0), (11, 2, 4, 2), (2, 3, -4, 0)\}$. Permuting the basis vectors gives an equivalent quadratic form:

$$x^2 - 3y^2 + 222z^2 + 11t^2 + 36yz - 20yt - 122zt.$$

Witt cancellation guarantees that $-3y^2 + 222z^2 + 11t^2 + 36yz - 20yt - 122zt$ is equivalent to $-y^2 + z^2 - 33t^2$ on $\mathbb{Q}^3$, which is not obvious because

$$\{(0, 2, 1, 0), (0, -1, 16, 1), (2, 3, -4, 0)\}$$

does not span the same 3-dimensional subspace of $\mathbb{Q}^4$ as

$$\{(0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}.$$

2.3 Applications

Using our previous work, we can classify non-degenerate quadratic forms up to equivalence over $\mathbb{C}$, $\mathbb{R}$, and finite fields $\mathbb{F}$ of odd characteristic.

Theorem 2.71. *Over $\mathbb{C}$, a non-degenerate quadratic form is determined up to equivalence by its dimension.*

Proof. Let Q be an n-dimensional non-degenerate quadratic form over $\mathbb{C}$. Then

$$Q(x_1e_1 + \cdots + x_ne_n) = x_1^2 + \cdots + x_n^2$$

in some orthogonal basis $\{e_1, \ldots, e_n\}$ since every non-zero number in $\mathbb{C}$ is a square. Thus, there is only one non-degenerate quadratic form up to equivalence in each dimension. □

Theorem 2.72. *Over $\mathbb{R}$, a non-degenerate quadratic form is determined up to equivalence by its dimension and the number of positive coefficients in a diagonal representation of the quadratic form.*

Proof. Let Q be an n-dimensional non-degenerate quadratic form over $\mathbb{R}$. Then

$$Q(x_1e_1 + \cdots + x_ne_n) = x_1^2 + \cdots + x_p^2 - (x_{p+1}^2 + \cdots + x_n^2) \tag{2.13}$$

in some orthogonal basis $\{e_1, \ldots, e_n\}$ since $\mathbb{R}^\times/\mathbb{R}^{\times^2}$ is represented by $\{1, -1\}$. We offer two proofs that Q determines p. For the first proof, suppose

$$Q(x_1e_1' + \cdots + x_ne_n') = x_1^2 + \cdots + x_{p'}^2 - (x_{p'+1}^2 + \cdots + x_n^2) \tag{2.14}$$

in a different basis $\{e_1', \ldots, e_n'\}$. Consider the subspaces U and U' spanned by $\{e_1, \ldots, e_p\}$ and $\{e_{p'+1}', \ldots, e_n'\}$ respectively. On $U - \{0\}$, Q takes only positive values whereas Q takes only negative values on $U' - \{0\}$. Thus $U \cap U' = \{0\}$, so

$$\dim(U + U') = \dim U + \dim U' = p + (n - p').$$

Since Q is n-dimensional, $\dim(U + U') \leq n$, so $p + n - p' \leq n$. This implies $p \leq p'$. In a similar way, $p' \leq p$. Thus $p = p'$. Hence p is independent of the choice of basis. This concludes the first proof.

For our second proof, we will induct on the dimension n and use Witt cancellation. Let $n = 1$. Since Q is non-degenerate either $Q(xe_1) = x^2$ or $Q(xe_1) = -x^2$ in a suitable basis $\{e_1\}$. So our theorem is trivially true.

Now suppose that the theorem is true in dimension $n-1$, where $n \geq 2$. Let Q be a non-degenerate n-dimensional quadratic form. Write Q in the form (2.13) and (2.14) in orthogonal bases $\{e_1, \ldots, e_n\}$ and $\{e_1', \ldots, e_n'\}$. From these expressions we have that $p > 0$ if and only if Q takes on some positive value and similarly for p'. So Q takes on no positive values if and only if $p = 0$ if and only if $p' = 0$.

Now assume that Q takes on positive values. This means $p > 0$ and $p' > 0$. So

$$\mathbb{R}e_1 \perp (\mathbb{R}e_2 \perp \cdots \perp \mathbb{R}e_n) = \mathbb{R}e_1' \perp (\mathbb{R}e_2' \perp \cdots \perp \mathbb{R}e_n')$$

and $\mathbb{R}e_1 \cong \mathbb{R}e_1'$. Thus Witt cancellation gives us

$$\mathbb{R}e_2 \perp \cdots \perp \mathbb{R}e_n \cong \mathbb{R}e_2' \perp \cdots \perp \mathbb{R}e_n'.$$

These $(n-1)$-dimensional quadratic spaces are non-degenerate by Corollary 2.23, so by induction $p - 1 = p' - 1$, and hence $p = p'$. □

Remark 2.73. Theorem 2.72 shows that there are $n+1$ equivalence classes of non-degenerate n-dimensional quadratic forms over $\mathbb{R}$. Furthermore, knowing any two values of the dimension n, the number of positive coefficients p or the number of negative coefficients q gives the third value since they are related by the equation $n = p + q$.

We now turn to quadratic forms over a finite field $\mathbb{F}$ of odd characteristic.

Lemma 2.74. *Every non-degenerate quadratic form over $\mathbb{F}$ with dimension ≥ 2 is universal.*

Proof. Let $q = \#\mathbb{F}$. For the 2-dimensional case, let $Q(xe_1 + ye_2) = ax^2 + by^2$ for $a, b \in \mathbb{F}$ in some orthogonal basis $\{e_1, e_2\}$. Since Q is non-degenerate, a and b are non-zero. For any $c \in \mathbb{F}$, consider the equation

$$ax^2 = c - by^2. \tag{2.15}$$

Both the left side and the right side of (2.15) take on $(q+1)/2$ values since $a, b \neq 0$. Thus there is an overlap in the values that the two sides take on. So for some $x_0, y_0 \in \mathbb{F}$ we have $ax_0^2 = c - by_0^2$, so $ax_0^2 + by_0^2 = c$.

For dimension $n \geq 2$, let $Q(x_1e_1 + \cdots + x_ne_n) = a_1x_1^2 + \cdots + a_nx_n^2$ in an orthogonal basis $\{e_1, \ldots, e_n\}$ where $a_i \in \mathbb{F}^\times$. Set $n-2$ of the variables equal to zero, so we are reduced to the 2-dimensional case. □

Lemma 2.75. *Every non-degenerate n-dimensional quadratic form over $\mathbb{F}$ can be written in some basis as $x_1^2 + \cdots + x_{n-1}^2 + dx_n^2$ for some $d \in \mathbb{F}^\times$.*

Proof. We will induct on the dimension. Let $n = 1$. Then $Q(x) = dx^2$ with $d \in \mathbb{F}^\times$ since Q is non-degenerate. Now let $n \geq 2$. By Lemma 2.74, Q represents 1 so

$$Q(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2$$

in some orthogonal basis $\{e_1, e_2, \ldots, e_n\}$. Let $U = \operatorname{span}\{e_2, \ldots, e_n\} = e_1^\perp$, so U is non-degenerate. Restricting Q to U gives a non-degenerate $(n-1)$-dimensional quadratic form. By induction there is a basis $\{e_2', \ldots, e_n'\}$ of U such that

$$Q(x_2e_2' + \cdots + x_ne_n') = x_2^2 + \cdots + x_{n-1}^2 + dx_n^2$$

for some $d \in \mathbb{F}^\times$. Attaching e_1 to this basis gives

$$Q(x_1e_1 + x_2e_2' + \cdots + x_ne_n') = x_1^2 + \cdots + x_{n-1}^2 + dx_n^2.$$

□

Theorem 2.76. *Over a finite field of odd characteristic, a non-degenerate quadratic form is determined up to equivalence by its dimension and its discriminant.*

Proof. By Lemmas 2.74 and 2.75, any n-dimensional non-degenerate quadratic form over a finite field $\mathbb{F}$ can be written in some orthogonal basis as

$$Q(x_1e_1 + \cdots + x_ne_n) = x_1^2 + \cdots + x_{n-1}^2 + dx_n^2$$

where $d \in \mathbb{F}^\times$. Note that $\operatorname{disc}(Q) = d$. Since $\mathbb{F}^\times/\mathbb{F}^{\times 2}$ has size 2, there are two quadratic forms for each dimension up to equivalence, one where $\operatorname{disc}(Q)$ is a square and one where $\operatorname{disc}(Q)$ is not a square. □

Chapter 3

Local Fields

Although the only context in which we will use local fields in our proof of the Hasse–Minkowski theorem is as completions of $\mathbb{Q}$ and $\mathbb{F}(T)$, we treat them in this chapter independently of their origin. That is, we do not care so much as to how they came into being (as a metric completion), we just care about their general characteristics and structure. The chapter is designed in this way because one can contemplate the Hasse–Minkowski theorem over any global field, where general local fields will play the role of the p-adic fields in the Hasse–Minkowski theorem over $\mathbb{Q}$. For a review of p-adic fields see [3] or [5].

3.1 Generalities

Our goal in this section is to take what the reader should already know about the fields $\mathbb{Q}_p$ and examine analogous results in a more general context. Most of the results will be stated without proof since the p-adic proofs apply without significant change. In this section $\mathbb{F}$ denotes any finite field (possibly of characteristic 2).

In this section K is a complete field with respect to a non-archimedean absolute value $|\cdot|$. The *integer ring* of K is $\mathcal{O} = \{x \in K : |x| \leq 1\}$, whose unit group is $\mathcal{O}^\times = \{x \in K : |x| = 1\}$. The complement of $\mathcal{O}^\times$ in $\mathcal{O}$ is $\mathfrak{m} = \{x \in K : |x| < 1\}$ which is a maximal ideal. The field $\mathcal{O}/\mathfrak{m}$ is called the *residue field* of K.

Example 3.1. If $K = \mathbb{Q}_p$ then $\mathcal{O} = \mathbb{Z}_p$ and $\mathfrak{m} = p\mathbb{Z}_p$. Furthermore,

$$\mathcal{O}/\mathfrak{m} = \mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{Z}/p\mathbb{Z}.$$

Example 3.2. Let $K = F((X))$, the field of formal Laurent series in X over any field F. Set $\mathrm{ord}(\sum a_n X^n)$ to be the least n such that $a_n \neq 0$ and $\mathrm{ord}(0) = \infty$. Choose $c \in \mathbb{R}$ so that $c > 1$ and set $|f| = c^{-\mathrm{ord}(f)}$ for $f \in F((X))$. Then $\mathcal{O} = F[[X]]$, $\mathfrak{m} = (X)$, and $\mathcal{O}/\mathfrak{m} = F[[X]]/(X)$, which is isomorphic to F.

Recall that there is an absolute value on $\mathbb{Q}$ corresponding to each prime p: for $r \in \mathbb{Q}$,

$$|r|_p = \begin{cases} p^{-\mathrm{ord}_p(r)}, & \text{for } r \in \mathbb{Q}^\times, \\ 0, & \text{for } r = 0, \end{cases} \tag{3.1}$$

where $\mathrm{ord}_p(r)$ is the power of p in r. Completing $\mathbb{Q}$ with respect to these absolute values yields the fields $\mathbb{Q}_p$ while completing $\mathbb{Q}$ with respect to the usual absolute value

$$|r|_\infty = \begin{cases} r, & \text{for } r \geq 0, \\ -r, & \text{for } r \leq 0, \end{cases} \tag{3.2}$$

gives $\mathbb{R}$.

Similarly, letting $\mathbb{F}$ be a finite field and $q := \#\mathbb{F}$, there is an absolute value on $\mathbb{F}[T]$ corresponding to each monic irreducible π in $\mathbb{F}(T)$. These absolute values are defined as follows: for $f \in \mathbb{F}(T)$,

$$|f|_\pi = \begin{cases} (q^{\deg \pi})^{-\mathrm{ord}_\pi(f)}, & \text{for } f \in \mathbb{F}(T)^\times, \\ 0, & \text{for } f = 0, \end{cases} \tag{3.3}$$

where $\mathrm{ord}_\pi(f)$ is the power of π in f. Also, there is another absolute value on $\mathbb{F}(T)$: for $f \in \mathbb{F}(T)$,

$$|f|_\infty = \begin{cases} q^{\deg f}, & \text{for } f \in \mathbb{F}(T)^\times, \\ 0, & \text{for } f = 0. \end{cases} \tag{3.4}$$

When we complete $\mathbb{F}(T)$ with respect to these absolute values we get the fields denoted by $\mathbb{F}(T)_\pi$ and $\mathbb{F}(T)_\infty$. Although the absolute value $|\cdot|_\infty$ seems different, it is really the same as the other absolute values on $\mathbb{F}(T)$ in the sense that they are all non-archimedean. The key is to realize that the degree function is just the negative of the power of $1/T$ in an element of $\mathbb{F}(T)$. That is, if $f = (1/T)^m u$ where $\deg u = 0$, then $\deg f = -m$. Another important thing to come to terms with is that the ring $\mathbb{F}[T]$ is *not* contained in the integer ring of $\mathbb{F}(T)_\infty$. Just look at the absolute value: $|f|_\infty > 1$ for all non-constant $f \in \mathbb{F}[T]$, so $|f - g|_\infty > 1$ when f and g are in $\mathbb{F}[T]$ and $f - g$ is non-constant. This means $\mathbb{F}[T]$ is a discrete subset of $\mathbb{F}(T)_\infty$.

Remark 3.3. As a convention we use v as a label for a general non-trivial absolute value on $\mathbb{Q}$ or $\mathbb{F}(T)$ as in (3.1), (3.2), (3.3), or (3.4). The notation $\mathbb{Q}_v$ (where $\mathbb{Q}_\infty = \mathbb{R}$ and $|\ |_\infty$ denotes the usual archimedean absolute value on $\mathbb{Q}$) and $\mathbb{F}(T)_v$ is used to represent any completion of $\mathbb{Q}$ and $\mathbb{F}(T)$. In this way we can write all of these completions in a common form.

The absolute values of $\mathbb{Q}$ and $\mathbb{F}(T)$ that we described are a complete list of non-trivial absolute values up to equivalence in the following sense.

Definition 3.4. Two absolute values on a field are *equivalent* when they define the same topology.

Theorem 3.5. *If $|\ |_1$ and $|\ |_2$ are equivalent absolute values on a field then $|\ |_1 = |\ |_2^t$ for some $t > 0$. In particular, each of the properties $|x| < 1$, $|x| = 1$, $|x| > 1$ does not change if the absolute value $|\ |$ is replaced by an equivalent absolute value.*

Proof. See [4, Theorem 9.1]. □

Theorem 3.6 (Ostrowski)**.** *Any archimedean absolute value on $\mathbb{Q}$ is equivalent to $|\ |_\infty$ and any non-trivial non-archimedean absolute value on $\mathbb{Q}$ is equivalent to $|\cdot|_p$ for a unique prime p. Similarly, any non-trivial absolute value on $\mathbb{F}(T)$ is equivalent to $|\cdot|_\infty$ or to $|\cdot|_\pi$ for one monic irreducible π in $\mathbb{F}[T]$.*

Proof. See [4, §9.3], which includes the classification of absolute values on $F(T)$ that are trivial on F, where F is any field. Note that absolute values on $\mathbb{F}(T)$ are automatically trivial on $\mathbb{F}$ since every element of $\mathbb{F}^\times$ is a root of unity. □

The next two results link together all the absolute values on each of $\mathbb{Q}$ and $\mathbb{F}(T)$ as in (3.1), (3.2), (3.3), and (3.4).

Theorem 3.7. *For $r \in \mathbb{Q}^\times$, $|r|_v \neq 1$ for at most finitely many v and*

$$\prod_v |r|_v = 1.$$

Proof. Write $r = \pm p_1^{e_1} \cdots p_r^{e_m}$ where the p_i are distinct primes. Then

$$\prod_{p_i} |r|_{p_i} = p_1^{-e_1} \cdots p_m^{-e_m},$$

$|r|_p = 1$ for $p \neq p_i$, and $|r|_\infty = p_1^{e_1} \cdots p_m^{e_m}$, so

$$\prod_v |r|_v = 1.$$

□

Theorem 3.8. *For $f \in \mathbb{F}(T)^\times$, $|f|_v \neq 1$ for at most finitely many v and*

$$\prod_v |f|_v = 1.$$

Proof. Let $q = \#\mathbb{F}$. Write $f = c\pi_1^{e_1} \cdots \pi_m^{e_m}$ for distinct monic irreducibles π_i and $c \in \mathbb{F}^\times$. This means $\prod_{\pi_i} |f|_{\pi_i} = q^{-e_1 \deg \pi_1} \cdots q^{-e_m \deg \pi_m}$, $|f|_\pi = 1$ for $\pi \neq \pi_i$, and $|f|_\infty = q^{\deg f}$. Since $\deg f = e_1 \deg \pi_1 + \cdots + e_m \deg \pi_m$,

$$\prod_v |f|_v = 1.$$

□

We return to the general scenario. The following theorem will play an important role in the Hasse–Minkowski theorem for dimension at least 5.

Theorem 3.9 (Approximation Theorem)**.** *Let $|\ |_1, \ldots, |\ |_r$ be inequivalent non-trivial absolute values on a field F, a_k an element of the completion of F with respect to $|\ |_k$ and $\varepsilon \in \mathbb{R}$ with $\varepsilon > 0$. Then there exists an $a \in F$ such that*

$$|a - a_k|_k < \varepsilon \text{ for all } 1 \leq k \leq r.$$

Proof. See [4, §9.2]. □

Over fields with a non-trivial non-archimedean absolute value there is an analogue of Newton's method for finding reals roots of polynomials; this is the next theorem.

Theorem 3.10 (Hensel's Lemma)**.** *Let $f(X) = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ be a polynomial with coefficients in $\mathcal{O}$. Suppose that there exists $\alpha_0 \in \mathcal{O}$ such that:*

$$|f(\alpha_0)| < |f'(\alpha_0)|^2,$$

where f' is the (formal) derivative of f. Then there exists an $\alpha \in \mathcal{O}$ such that $\alpha \equiv \alpha_0 \bmod \mathfrak{m}$ and $f(\alpha) = 0$.

Proof. See [6, Proposition 7.6, Chapter XII]. □

Example 3.11. If the residue field is not of characteristic 2 then $a \in \mathcal{O}^\times$ is a square in K if and only if a is a square in $\mathcal{O}/\mathfrak{m}$. Consider the polynomial $X^2 - a$ and apply Hensel's lemma where $\alpha_0^2 \equiv a \bmod \mathfrak{m}$. Note that 2 is in $\mathcal{O}^\times$.

Example 3.12. If $\mathcal{O}/\mathfrak{m}$ has characteristic 2 and K has characteristic 0 (e.g., $K = \mathbb{Q}_2$) then $a \in \mathcal{O}^\times$ is a square in K if and only if a is a square modulo $4\mathfrak{m}$. Again consider the polynomial $f(X) = X^2 - a$. For $\alpha_0 \in \mathcal{O}^\times$, $|f'(\alpha_0)|^2 = |4|$. So Hensel's lemma applies when a is congruent to a square modulo $4\mathfrak{m}$.

The next result is a multivariable version of Hensel's lemma.

Theorem 3.13. *Let $f(X_1, \ldots, X_n)$ be a polynomial with coefficients in $\mathcal{O}$. Suppose there are $\gamma_1, \ldots, \gamma_n$ in $\mathcal{O}$ such that for some i*

$$|g(\gamma_1, \ldots, \gamma_n)| < \left| \frac{\partial g}{\partial X_i}(\gamma_1, \ldots, \gamma_n) \right|^2.$$

Then there is $\alpha \in \mathcal{O}$ such that $g(\gamma_1, \ldots, \alpha, \ldots, \gamma_n) = 0$.

Proof. Let $f(X) = g(\gamma_1, \ldots, \gamma_{i-1}, X, \gamma_{i+1}, \ldots, \gamma_n)$. Then

$$f'(X) = \frac{\partial g}{\partial X_i}(\gamma_1, \ldots, \gamma_{i-1}, X, \gamma_{i+1}, \ldots, \gamma_n).$$

Applying Hensel's lemma to f with $\alpha_0 = \gamma_i$ gives a root of g. □

Definition 3.14. When $\mathfrak{m}$ is principal, a *uniformizer* of K is any element π that generates $\mathfrak{m}$: $\mathfrak{m} = (\pi) = \pi\mathcal{O}$.

Remark 3.15. When K is discretely valued and $|K^\times| = c^{\mathbb{Z}}$ with $c > 1$, a uniformizer π can be equivalently defined as any element of largest absolute value less than one: $|\pi| = 1/c$.

Example 3.16. In $\mathbb{Q}_p$, the normal choice of uniformizer is p. However, one can choose any element with absolute value $1/p$; for example, $-p$ works too.

Example 3.17. In $F((X))$ for any field F, the usual uniformizer is X. A different uniformizer is $X + X^2 = X(1 + X)$.

Proposition 3.18. *Let π be a uniformizer of K. Every $x \in K^\times$ is $\pi^m u$ for unique $m \in \mathbb{Z}$ and u in $\mathcal{O}^\times = \{x \in K : |x| = 1\}$. Moreover, any two uniformizers are equal up to a unit multiple and $x \in \mathcal{O}$ if and only if $m \geq 0$.*

Definition 3.19. Let π be a uniformizer of K. For $x \in K$, the *valuation* of x is the power of π in x. That is, if $x = \pi^m u$ for $u \in \mathcal{O}^\times$ then the valuation of x is m. This is independent of the choice of π. Denote the valuation of x by $\mathrm{ord}(x)$.

Using the valuation function and setting $c = |\pi|^{-1} > 1$, we have

$$|x| = |\pi|^m = \begin{cases} c^{-\mathrm{ord}(x)}, & \text{for } x \in K^\times, \\ 0, & \text{for } x = 0. \end{cases}$$

We are now ready to define a local field.

Definition 3.20. A *local field* is a field which is complete with respect to a discrete (non-archimedean) absolute value and which has a finite residue field.

So what do the elements of a local field K look like? Fixing a uniformizer π and a set S of representatives for $\mathcal{O}/\mathfrak{m}$ (using 0 to represent the class of 0), any x in $\mathcal{O}$ has a unique π-adic expansion:

$$x = c_0 + c_1\pi + c_2\pi^2 + \cdots \tag{3.5}$$

where $c_i \in S$ and the c_i's are uniquely determined by x (S is fixed!). What if $x \in K$, but x is not in $\mathcal{O}$? Then x can be written as $x = y/\pi^m$ where $m \geq 1$ and $y \in \mathcal{O}$. Writing y as a π-adic expansion and dividing by π^m gives

$$x = c_0\pi^{-m} + c_1\pi^{-m+1} + \cdots + c_{m-1}\pi^{-1} + c_m + c_{m+1}\pi + \cdots \tag{3.6}$$

where $c_i \in S$ and the c_i's are uniquely determined by x.

Using Example 3.11, we can classify squares in a local field. The next two examples discuss squares in $\mathbb{F}(T)_v$.

Example 3.21. For a non-zero $f \in \mathbb{F}(T)_\pi$, write $f = \pi^n g$ where $|g|_\pi = 1$. Then f is a square in $\mathbb{F}(T)_\pi$ if and only if n is even and g is a square modulo π. Indeed, if f is a square then n is even since π is a uniformizer in $\mathbb{F}(T)_\pi$. Thus we are reduced to showing that g is a square if and only if g is a square modulo π. This last statement was proved in Example 3.11.

Example 3.22. The field $\mathbb{F}(T)_\infty$ has a uniformizer $1/T$ and is isomorphic to $\mathbb{F}((1/T))$. A non-zero $f \in \mathbb{F}[T]$ is a square in $\mathbb{F}(T)_\infty$ if and only if $\deg f$ is even and $\operatorname{lead} f$ (the leading coefficient of f) is a square in $\mathbb{F}^\times$. To see why, write

$$\begin{aligned} f &= c_n T^n + \cdots + c_1 T + c_0 \\ &= T^n (c_n + c_{n-1}\frac{1}{T} + \cdots + c_0 \frac{1}{T^n}). \end{aligned}$$

Here $\operatorname{ord}(f) = -n = -\deg f$ and $\operatorname{lead} f = c_n$ Note that

$$c_n + c_{n-1}T^{-1} + \cdots + c_0 T^{-n}$$

is a unit and that the residue field of $\mathbb{F}(T)_\infty$ is isomorphic to $\mathbb{F}$. So if n is even and $c_n \in \mathbb{F}^{\times 2}$ then f is a square in $\mathbb{F}(T)_\infty$ by Hensel's lemma. Conversely, if f is a square in $\mathbb{F}(T)_\infty$ then its valuation $-n$ is even so $c_n + c_{n-1}T^{-1} + \cdots + c_0 T^{-n}$ must be a square in the units $\mathbb{F}[[1/T]]$ of $\mathbb{F}(T)_\infty$. Reduce modulo T^{-1} to see that c_n must be a square in $\mathbb{F}$.

Theorem 3.23. *Let K be a field complete with respect to a non-trivial non-archimedean absolute value. The following are equivalent:*

1) *K is locally compact (i.e., every point has a compact neighborhood),*

2) *$\mathcal{O}$ is compact,*

3) *K is a local field.*

Proof. (2) $\Rightarrow$ (1): Pick $a \in K$. Since $\mathcal{O}$ is a compact neighborhood of zero and the map $f(x) = a + x$ is a homeomorphism of K with itself, $a + \mathcal{O}$ is a compact neighborhood of a.

(1) $\Rightarrow$ (2): By assumption there is some compact neighborhood A of zero. Choose $\alpha \in \mathcal{O}$ with $0 < |\alpha| < 1$, so $\alpha^n \to 0$ as $n \to \infty$. Since A is a neighborhood of 0, for sufficiently large n the set $\alpha^n \mathcal{O} = \{x : |x| \leq |\alpha|^n\}$ is contained in A. Furthermore $\alpha^n \mathcal{O}$ is a closed subset of A so it is compact. Define $g : \alpha^n \mathcal{O} \to \mathcal{O}$ by $g(x) = x\alpha^{-n}$. Then g is a homeomorphism so $\mathcal{O}$ is compact.

(2) $\Rightarrow$ (3): The ideal $\mathfrak{m}$ is open in $\mathcal{O}$, so the cosets $x + \mathfrak{m}$ for $x \in \mathcal{O}$ form an open cover of $\mathcal{O}$. However, only a finite number of the cosets of $\mathfrak{m}$ are

needed to cover $\mathcal{O}$ since $\mathcal{O}$ is compact. Therefore, $\mathcal{O}/\mathfrak{m}$ is finite. To show $|K^\times|$ is discrete, suppose otherwise. Then there is a sequence in $\mathcal{O}$ such that

$$|x_1| < |x_2| < \cdots < 1$$

with $|x_i| \to 1$ as $i \to \infty$. Since $\mathcal{O}$ is compact, there is a convergent subsequence with a limit point, say x, such that $|x| = 1$. So if $|x - x_i| < 1$ then $|x_i| = 1$ by the non-archimedean absolute value, but $|x_i| < 1$ for all i. This is a contradiction.

(3) $\Rightarrow$ (2): We will show $\mathcal{O}$ is sequentially compact, which is the same as compactness since K is a metric space. Let S be a set of coset representatives of $\mathfrak{m}$ and π be a uniformizer. Let A_n be an infinite sequence in $\mathcal{O}$. From (3.5), an infinite number of terms of A_n must have some common c_0 as the first coefficient in their π-adic expansion since S is finite. Similarly, an infinite number of terms of A_n that have c_0 as the initial coefficient in their π-adic expansion also have some common c_1 as the second coefficient in their π-adic expansion since S is finite. Continuing in this manner ad infinitum we see that A_n contains a Cauchy sequence which converges to a limit point in $\mathcal{O}$ because K is complete and $\mathcal{O}$ is closed in K. □

Theorem 3.24. *Local fields are either finite field extensions of $\mathbb{Q}_p$ or fields $\mathbb{F}((X))$ with $\mathbb{F}$ finite.*

Proof. See [4, Theorem 9.16]. □

Remark 3.25. Note that in Theorem 3.24, $\mathbb{F}$ can have characteristic 2.

Remark 3.26. Finite extensions of $\mathbb{Q}_p$ are often referred to as *p-adic fields*. For example, a finite extension of $\mathbb{Q}_2$ is called a *2-adic field*.

One nice application of Hensel's lemma when K is a local field is that we can find a full set of $(q-1)$th roots of unity in K where q is the size of the residue field of K. To do this, consider the polynomial $f(X) = X^{q-1} - 1$. Then for each $a \in \mathcal{O}^\times$, $|f(a)| < 1$ while $|f'(a)| = 1$ so Hensel's lemma gives a root ζ of $f(X)$ such that $\zeta \equiv a \bmod \mathfrak{m}$. Thus each non-zero coset of $\mathfrak{m}$ has a root of f. Moreover, these must be all of the roots since

$$\deg f = q - 1 = \#(\mathcal{O}/\mathfrak{m})^\times.$$

Note that $\mu_{q-1} = \{\zeta : \zeta^{q-1} = 1\}$ is a subgroup of $\mathcal{O}^\times$. We now use this to analyze the structure of $\mathcal{O}^\times$.

Theorem 3.27. *For any local field K we have an isomorphism of groups $\mathcal{O}^\times \cong \mu_{q-1} \times (1+\mathfrak{m})$, where $q = \#(\mathcal{O}/\mathfrak{m})$.*

Proof. For $a \in \mathcal{O}^\times$, let $\zeta \in \mu_{q-1}$ such that $\zeta \equiv a \bmod \mathfrak{m}$. Let $u = a/\zeta$. So $u \in \mathcal{O}^\times$ and $u \equiv 1 \bmod \mathfrak{m}$. That is, $u \in 1+\mathfrak{m}$ and $a = \zeta u$. Thus $\mathcal{O}^\times = \mu_{q-1}(1+\mathfrak{m})$. Furthermore, μ_{q-1} and $1+\mathfrak{m}$ are subgroups of $\mathcal{O}^\times$ that intersect trivially. So $\mathcal{O}^\times \cong \mu_{q-1} \times (1+\mathfrak{m})$. □

Remark 3.28. Elements of $\mu_{q-1} \cup \{0\}$ are called *Teichmüller representatives*; they are sometimes the most natural set of representatives to use for $\mathcal{O}/\mathfrak{m}$.

Not only does Theorem 3.27 give structure to $\mathcal{O}^\times$, but it also shows that

$$K^\times \cong \pi^{\mathbb{Z}} \times \mu_{q-1} \times (1+\mathfrak{m}). \tag{3.7}$$

We conclude this section with an index inequality that will be useful in Section 3.3 (Theorem 3.63).

Theorem 3.29. *Let K be a local field with $\operatorname{char}(K) \neq 2$. Then*

$$[K^\times : K^{\times 2}] \geq 4.$$

Proof. Since $K^\times \cong \pi^{\mathbb{Z}} \times \mathcal{O}^\times$,

$$K^\times / K^{\times 2} \cong \mathbb{Z}/2\mathbb{Z} \times \mathcal{O}^\times / \mathcal{O}^{\times 2}.$$

So it is enough to show that $\mathcal{O}^\times \neq \mathcal{O}^{\times 2}$. If K has odd residue field characteristic use a non-square in μ_{q-1}. If K has a residue field of characteristic 2 then $\mu_{q-1} = \mu_{q-1}^2$ so we find a non-square in $1+\mathfrak{m}$. Set $u = 1+\pi$ where π is a uniformizer of K. Suppose u is square in $\mathcal{O}^\times$. Then $u = c^2$ for some $c \in \mathcal{O}^\times$. So $c \equiv 1 \bmod \pi$ since squaring is injective in characteristic 2 ($\operatorname{char}(\mathcal{O}/\mathfrak{m}) = 2$). Thus $c = 1 + d\pi$ for some $d \in \mathcal{O}$. Squaring gives

$$1 + \pi = 1 + 2d\pi + d^2\pi^2. \tag{3.8}$$

This is false, however, since the right side of (3.8) is $1 \bmod \pi^2$ because 2 is $0 \bmod \pi$ while the left side is $1 + \pi \bmod \pi^2$. Thus u is a non-square in $\mathcal{O}^\times$. □

3.2 Quadratic Forms Over Local Fields

We use the facts about local fields laid out in Section 3.1 and the discussion of quadratic forms from Chapter 1 to examine quadratic forms over local fields. For the whole of this section let K be a local field not of characteristic 2, $\mathcal{O}$ be its integer ring, and π be a uniformizer for K. The residue field characteristic of K might be 2.

Theorem 3.30. *Let Q be a non-degenerate quadratic form over K. Then, in some basis,*

$$Q(x_1, \ldots, x_n) = a_1 x_1^2 + \cdots + a_r x_r^2 + \pi(a_{r+1} x_{r+1}^2 + \cdots + a_n x_n^2) \tag{3.9}$$

where the a_i's are units.

Proof. We can assume that Q is diagonal so write

$$Q(x_1, \ldots, x_n) = a_1' x_1^2 + \cdots + a_n' x_n^2$$

with a_i' in $K^\times$. Then either $a_i' = \pi^{2e_i} a_i$ or $a_i' = \pi^{2e_i+1} a_i$ where a_i is a unit. By a linear change of variables (set $x_i' = \pi^{e_i} x_i$ and then rearrange the basis vectors accordingly) we have

$$Q(x_1, \ldots, x_n) = a_1 x_1^2 + \cdots + a_r x_r^2 + \pi(a_{r+1} x_{r+1}^2 + \cdots + a_n x_n^2)$$

just as we wanted. □

Any vector in K^n can be scaled so that it is a vector in $\mathcal{O}^n$. Those vectors in $\mathcal{O}^n$ which do not reduce to 0 in $(\mathcal{O}/\mathfrak{m})^n$ have a special name:

Definition 3.31. A vector $(\alpha_1, \ldots, \alpha_n)$ in $\mathcal{O}^n$ is called *primitive* if at least one of the α_i's is non-zero in the residue field.

Lemma 3.32. *If $v = (\alpha_1, \ldots, \alpha_n)$ in K^n and $v \neq 0$ then cv is primitive for some $c \in K^\times$.*

Proof. Let $\max_{1 \leq i \leq n} |\alpha_i| = |\alpha_{i_0}|$. Then $(1/\alpha_{i_0})v$ is primitive: $|\alpha_i/\alpha_{i_0}| \leq 1$ for all i and $\alpha_{i_0}/\alpha_{i_0} = 1$. □

Theorem 3.33. *Any quadratic form over any local field that has a null vector also has a primitive null vector.*

Proof. Let Q be a quadratic form over a local field K that has a null vector v. By Lemma 3.32, cv is primitive for some $c \in K^\times$ and

$$Q(cv) = c^2 Q(v) = 0.$$

□

Theorem 3.34. *When K has odd residue field characteristic, the equation*

$$a_1 x_1^2 + \cdots + a_r x_r^2 + \pi(a_{r+1} x_{r+1}^2 + \cdots + a_n x_n^2) = 0 \tag{3.10}$$

with all $a_i \in \mathcal{O}^\times$ has a non-trivial solution over K if and only if at least one of

$$a_1 x_1^2 + \cdots + a_r x_r^2$$

and

$$a_{r+1} x_{r+1}^2 + \cdots + a_n x_n^2$$

has a non-trivial solution over K.

Proof. We only prove the "only if" direction since the "if" direction is trivial. Let

$$F_1(x_1, \ldots, x_r) = a_1 x_1^2 + \cdots + a_r x_r^2$$

and

$$F_2(x_{r+1}, \ldots, x_n) = a_{r+1} x_{r+1}^2 + \cdots + a_n x_n^2.$$

Assume there is a non-trivial solution over K to (3.10), say $(\alpha_1, \ldots, \alpha_n)$. By Theorem 3.33, we may assume that every α_i is in $\mathcal{O}$ and that at least one of them is in $\mathcal{O}^\times$. Suppose $\alpha_i \not\equiv 0 \bmod \pi$ for some $i \leq r$. Then

$$F_1(\alpha_1, \ldots, \alpha_r) \equiv 0 \bmod \pi$$

and

$$\frac{\partial F_1}{\partial x_i}(\alpha_1, \ldots, \alpha_r) = 2a_i \alpha_i \not\equiv 0 \bmod \pi.$$

So there is a non-trivial solution to $F_1(x_1, \ldots, x_r) = 0$ over K by Theorem 3.13.

On the other hand, if $\alpha_1, \ldots, \alpha_r$ are all divisible by π, then at least one of $\alpha_{r+1}, \ldots, \alpha_n$ must be in $\mathcal{O}^\times$. Furthermore, $F_1(\alpha_1, \ldots, \alpha_r) \equiv 0 \bmod \pi^2$ so we can divide the congruence

$$a_1 \alpha_1^2 + \cdots + a_r \alpha_r^2 + \pi(a_{r+1} \alpha_{r+1}^2 + \cdots + a_n \alpha_n^2) \equiv 0 \bmod \pi^2$$

by π to see that

$$F_2(\alpha_{r+1}, \ldots, \alpha_n) \equiv 0 \bmod \pi.$$

Assume, without loss of generality, that $\alpha_{r+1} \in \mathcal{O}^\times$. Then

$$\frac{\partial F_2}{\partial x_{r+1}}(\alpha_{r+1}, \ldots, \alpha_n) \not\equiv 0 \bmod \pi$$

so applying Theorem 3.13 again gives a non-trivial solution to

$$F_2(x_{r+1}, \ldots, x_n) = 0$$

over K. □

Corollary 3.35. *When K has odd residue field characteristic, the diagonal quadratic form in* (3.9) *has a null vector if and only if the equation* (3.10) *has a primitive solution modulo π^2.*

Proof. This falls out of the proof of Theorem 3.34. □

Remark 3.36. Corollary 3.35 gives a finite check for deciding when a non-degenerate quadratic form over K has a null vector since $\mathcal{O}/\pi^2\mathcal{O}$ is finite.

Theorem 3.37. *When K has odd residue field characteristic and α, β, γ are in $\mathcal{O}^\times$, the quadratic form $\alpha x^2 + \beta y^2 + \gamma z^2$ has a null vector in K^3.*

Proof. Let q denote the size of $\mathcal{O}/\mathfrak{m}$, so q is odd. Consider the congruence

$$\alpha x^2 = -\beta y^2 - \gamma z^2 \bmod \pi.$$

Taking $z = 1$, both sides of the congruence take on $\frac{q+1}{2}$ values as x and y run over $\mathcal{O}/\mathfrak{m}$. Thus, for some x_0, y_0 in $\mathcal{O}/\mathfrak{m}$, the two sides take on a common value. Furthermore, π does not divide both x_0 and y_0 since $\gamma \not\equiv 0 \bmod \pi$. Without loss of generality assume $x_0 \not\equiv 0 \bmod \pi$. Then

$$x_0^2 \equiv \frac{-\beta}{\alpha} y_0^2 - \frac{\gamma}{\alpha} \bmod \pi$$

since α is invertible modulo π. So by Hensel's lemma there is a root $\tilde{x}_0$ to the polynomial $X^2 - (-\frac{\beta}{\alpha} y_0^2 - \frac{\gamma}{\alpha})$. Hence $(\tilde{x}_0, y_0, 1)$ is a null vector for $\alpha x^2 + \beta y^2 + \gamma z^2$. □

Example 3.38. The quadratic form $x^2 + y^2 + z^2$ has a null vector over any local field with odd residue field characteristic by Theorem 3.37.

Corollary 3.39. *When K has odd residue field characteristic, any quadratic form over K with dimension at least 5 has a null vector.*

Proof. By Theorem 2.32 it suffices to focus on non-degenerate quadratic forms Q. Write $Q = Q_1 + \pi Q_2$ where Q_1 and Q_2 are diagonal with coefficients in $\mathcal{O}^\times$ (Theorem 3.30). Either Q_1 or Q_2 is at least 3-dimensional since $\dim Q_1 + \dim Q_2 = \dim Q \geq 5$. For the Q_i that is at least 3-dimensional, set all but 3 variables equal to 0. Then use Theorem 3.37 to get a null vector for this Q_i. □

Remark 3.40. We will see later (Theorem 3.63) that Corollary 3.39 is also true when K is a 2-adic field.

Corollary 3.35 gives a necessary and sufficient condition for quadratic forms over a local field with odd residue field characteristic to have a null vector. Over local fields of characteristic 0 that have a residue field of characteristic 2 (i.e., 2-adic fields) the theory is a little more subtle. The following theorem gives a necessary and sufficient condition for deciding when a quadratic form over a 2-adic field has a null vector.

Theorem 3.41. *Let K be a 2-adic field. The quadratic form Q in (3.9) has a null vector if and only if the congruence $Q \equiv 0 \bmod 4\pi^2$ has a primitive solution.*

Proof. Let $\dim Q = n$. Write

$$Q(x_1, \ldots, x_n) = a_1 x_1^2 + \cdots + a_r x_r^2 + \pi(a_{r+1} x_{r+1}^2 + \cdots + a_n x_n^2)$$

with $a_i \in \mathcal{O}^\times$.

The "only if" direction of this theorem is clear from Theorem 3.33; we prove the "if" direction. Suppose $Q(\alpha_1, \ldots, \alpha_n) \equiv 0 \bmod 4\pi^2$ with $\alpha_i \in \mathcal{O}$ where at least one α_i is a unit. Assume that $\alpha_i \not\equiv 0 \bmod \pi$ for some $i \leq r$. This implies

$$\frac{\partial Q}{\partial x_i}(\alpha_1, \ldots, \alpha_n) = 2a_i\alpha_i \not\equiv 0 \bmod 2\pi.$$

Thus

$$|Q(\alpha_1, \ldots, \alpha_n)| \leq |4\pi^2| < |4| = \left|\frac{\partial Q}{\partial x_i}(\alpha_1, \ldots, \alpha_n)\right|^2,$$

so Q has a null vector by Theorem 3.13. Note that we only needed the congruence $Q(\alpha_1, \ldots, \alpha_n) \equiv 0 \bmod 4\pi$, not modulo $4\pi^2$.

Now assume that α_i is divisible by π for $1 \leq i \leq r$ and write $\alpha_i = \pi\eta_i$ for these i. So by our assumption

$$\pi^2 \sum_{i=1}^{r} a_i \eta_i^2 + \pi \sum_{i=r+1}^{n} a_i \alpha_i^2 \equiv 0 \bmod 4\pi^2 \tag{3.11}$$

where at least one of $\alpha_{r+1}, \ldots, \alpha_n$ is a unit. Dividing (3.11) by π gives

$$\sum_{i=r+1}^{n} a_i \alpha_i^2 + \pi \sum_{i=1}^{r} a_i \eta_i^2 \equiv 0 \bmod 4\pi. \tag{3.12}$$

Using (3.12) and the same reasoning as before, the quadratic form

$$\frac{1}{\pi} Q(\pi x_1, \ldots, \pi x_r, x_{r+1}, \ldots, x_n)$$

has a null vector. Then trivially Q has a null vector. □

Corollary 3.42. *For any 2-adic field K, the equation $a_1 x_1^2 + \cdots + a_n x_n^2 = 0$ where each a_i is in $\mathcal{O}^\times$ has a non-trivial solution in K if and only if the congruence*

$$a_1 x_1^2 + \cdots + a_n x_n^2 \equiv 0 \bmod 4\pi$$

has a primitive solution.

Proof. This corollary falls out of the proof of Theorem 3.41. □

Example 3.43. Recall the quadratic form $x^2 + y^2 + z^2$ from Example 3.38. Consider the congruence in $\mathbb{Z}_2$

$$x^2 + y^2 + z^2 \equiv 0 \bmod 8\mathbb{Z}_2.$$

There are no primitive solutions to this congruence so $x^2 + y^2 + z^2$ has no null vectors in $\mathbb{Q}_2^3$.

3.3 The Hilbert Symbol

In this section we define the Hilbert symbol over a local field and discuss some of its properties and applications. In Chapters 4 and 5 it will play an important part in the proof of the Hasse–Minkowski theorem over $\mathbb{Q}$ and $\mathbb{F}(T)$. For this section K is a local field *not of characteristic* 2 (note that K could have a residue field of characteristic 2, i.e., K could be a 2-adic field).

Definition 3.44. Let K be a local field not of characteristic 2. For $a, b \in K^\times$, their *Hilbert symbol* is

$$(a,b)_K = \begin{cases} 1, & \text{if } b = x^2 - ay^2 \text{ for some } x, y \in K, \\ -1, & \text{otherwise.} \end{cases}$$

Example 3.45. For a, b, c in $K^\times$ we have

- $(a, -a)_K = 1$,
- $(1, b)_K = 1$,
- $(a, bc^2)_K = (ac^2, b)_K = (a, b)_K$.

Lemma 3.46. *Let F be any field. For $a \in F^\times$,*

$$\{x^2 - ay^2 \neq 0 : x, y \in F\}$$

is a subgroup of $F^\times$.

Proof. If a is a square in F, then $x^2 - ay^2$ is universal by Theorem 2.38. So $\{x^2 - ay^2 \neq 0 : x, y \in F\}$ is $F^\times$.

If a is not a square in F, then $F(\sqrt{a})$ is a quadratic extension of F. For any $\alpha \in F(\sqrt{a})$, write $\alpha = x + y\sqrt{a}$. So the norm $\mathrm{N}_{F(\sqrt{a})/F}(\alpha)$ equals $x^2 - ay^2$. In particular, $\{x^2 - ay^2 \neq 0 : x, y \in F\}$ is the image of the norm function on $F(\sqrt{a})^\times$. Thus it is a subgroup of $F^\times$. □

Theorem 3.47. *Over a field F not of characteristic 2, with a and b in $F^\times$, the following are equivalent:*

1) *$x_1^2 - ax_2^2 - bx_3^2 + abx_4^2$ has a null vector over F,*
2) *$b = x^2 - ay^2$ for some $x, y \in F$,*
3) *$ax^2 + by^2 - z^2$ has a null vector over F.*

Proof. If a is a square in F then (1) and (3) are trivially true and (2) is true from Theorem 2.38. So let a not be a square in F. Suppose

$$x_1^2 - ax_2^2 - bx_3^2 + abx_4^2 = 0$$

non-trivially. Then

$$x_1^2 - ax_2^2 = b(x_3^2 - ax_4^2)$$

where $x_1^2 - ax_2^2$ and $x_3^2 - ax_4^2$ are non-zero since a is not a square in F. Thus $b = x^2 - ay^2$ for some $x, y \in F$ because $\{x^2 - ay^2 \neq 0 : x, y \in F\}$ is a subgroup of $F^\times$ (Lemma 3.46).

Now suppose $b = x^2 - ay^2$ for some $x, y \in F$. Then relabeling the variables, $ax^2 + by^2 - z^2$ has a null vector over F (let $y = 1$).

If $ax^2 + by^2 - z^2$ has a null vector (x_0, y_0, z_0) over F then it is clear that $x_1^2 - ax_2^2 - bx_3^2 + abx_4^2$ does too: $x_1 = z_0$, $x_2 = x_0$, $x_3 = y_0$, and $x_4 = 0$. □

The algebraic properties of the Hilbert symbol will follow from the next result. It is the key fact from which the theory of quadratic forms over local fields is built.

Theorem 3.48. *Let L be a quadratic extension of a local field K not of characteristic 2. Then $[K^\times : \mathrm{N}_{L/K}(L^\times)] = 2$.*

A proof of Theorem 3.48 is given in Appendix B.

Remark 3.49. Theorem 3.48 is also true when L/K is a quadratic Galois extension of a local field of characteristic 2.

Corollary 3.50. *Let K be a local field not of characteristic 2. The Hilbert symbol on K has the following properties:*

1) $(a, b)_K = (b, a)_K$,

2) *if a is not a square in K, then there is some b such that $(a, b)_K = -1$,*

3) $(a, bb')_K = (a, b)_K(a, b')_K$ *and* $(aa', b)_K = (a, b)_K(a', b)_K$.

Proof. Throughout the proof let $a, b \in K^\times$.

(1) By Theorem 3.47, $(a, b)_K = 1$ exactly when $ax^2 + by^2 - z^2 = 0$ is non-trivially solvable over K. Similarly, $(b, a)_K = 1$ when $bx^2 + ay^2 - z^2 = 0$ is non-trivially solvable over K. Clearly $ax^2 + by^2 - z^2$ and $bx^2 + ay^2 - z^2$ are equivalent quadratic forms (just permute the basis vectors of one of the forms to get the other) so $(a, b)_K = 1$ if and only if $(b, a)_K = 1$. Thus $(a, b)_K = (b, a)_K$.

(2) If a is not a square in K, then $L = K(\sqrt{a})$ is a quadratic extension of K so $\mathrm{N}_{L/K}(x + y\sqrt{a}) = x^2 - ay^2$. By Theorem 3.48, there a $b \notin \mathrm{N}_{L/K}(L^\times)$. So $(a, b)_K \neq 1$.

(3) By (1) it is enough to show that $(a, bb')_K = (a, b)_K(a, b')_K$. Suppose a is a square. Then $x^2 - ay^2$ is universal. So the equation is trivially true since $(a, b)_K = 1$ for all $b \in K^\times$.

If a is not a square, let $L = K(\sqrt{a})$. Suppose $(a,b)_K = 1$ and $(a,b')_K = 1$. Then $b \in \mathrm{N}_{L/K}(L^\times)$ and $b' \in \mathrm{N}_{L/K}(L^\times)$. So bb' is in $\mathrm{N}_{L/K}(L^\times)$ since $\mathrm{N}_{L/K}(L^\times)$ is a group. Thus $(a,bb')_K = 1$.

Suppose exactly one of $(a,b)_K, (a,b')_K$ is -1. Without loss of generality assume $(a,b)_K = -1$ and $(a,b')_K = 1$. Then $b \notin \mathrm{N}_{L/K}(L^\times)$ and $b' \in \mathrm{N}_{L/K}(L^\times)$. Suppose $bb' \in \mathrm{N}_{L/K}(L^\times)$. Then $b = bb'/b'$ is in $\mathrm{N}_{L/K}(L^\times)$ which is a contradiction. So $(a,bb')_K = -1$.

Suppose $(a,b)_K = -1$ and $(a,b') = -1$. Then $bb' \in \mathrm{N}_{L/K}(L^\times)$ since $\mathrm{N}_{L/K}(L^\times)$ has index 2 in $K^\times$ by Theorem 3.48. So $(a,bb')_K = 1$. □

Since $[\mathbb{R}^\times : \mathrm{N}_{\mathbb{C}/\mathbb{R}}(\mathbb{C}^\times)] = [\mathbb{R}^\times : \mathbb{R}^{\times 2}] = 2$ we can let $K = \mathbb{R}$ in Definition 3.44 (even though $\mathbb{R}$ is not a local field) and the resulting symbol $(\cdot,\cdot)_{\mathbb{R}}$ will have the same algebraic properties as the Hilbert symbols over local fields from Corollary 3.50. Denote the Hilbert symbol over $\mathbb{R}$ by $(a,b)_{\mathbb{R}}$ for $a, b \in \mathbb{R}^\times$.

Example 3.51. The Hilbert symbol $(-1,-1)_K$ is 1 for K a local field of odd residue field characteristic (Theorem 3.37). When $K = \mathbb{R}$ or $K = \mathbb{Q}_2$, however, $(-1,-1)_K$ is -1 (look at the equation $-x^2 - y^2 - z^2 = 0$ and recall Example 3.43). By Corollary 3.42, $(-1,-1)_K = 1$ when K is a 2-adic field if and only if $x^2 + y^2 + z^2 \equiv 0 \bmod 4\pi$ has a primitive solution.

Example 3.52. The algebraic properties of the Hilbert symbol show that $(a,1)_K = (1,a)_K = 1$ and, more importantly for later calculations,

$$\begin{aligned}(a,a)_K &= (a,-a)_K(a,-1)_K \\ &= (a,-1)_K\end{aligned}$$

for any local field K ($\mathrm{char}(K) \neq 2$) and for $K = \mathbb{R}$.

This next result offers a formula for computing Hilbert symbols over local fields with odd residue field characteristic.

Theorem 3.53. *Let K be a local field with odd residue field characteristic and π a uniformizer. For $a, b \in K^\times$, write $a = \pi^m\varepsilon$ and $b = \pi^n\delta$ for m, n in $\mathbb{Z}$ and ε, δ in $\mathcal{O}^\times$. Then*

$$(a,b)_K = (-1)^{mn(q-1)/2}\chi(\bar{\varepsilon})^n\chi(\bar{\delta})^m$$

where q is the size of the residue field and $\chi : (\mathcal{O}_K/\mathfrak{m}_K)^\times \to \{\pm 1\}$ is the quadratic character on the residue field.

Proof. Denote the right side of the formula as $\langle a,b\rangle_K$.We first verify that $\langle a,b\rangle_K$, as a function of a and b, has the first and third algebraic properties of the Hilbert symbol in Corollary 3.50. The symmetry is obvious so we just check that $\langle a,b\rangle_K$ is multiplicative in a. Let a,a',b be in $K^\times$. Write $a=\pi^m\varepsilon$, $a'=\pi^{m'}\varepsilon'$ and $b=\pi^n\delta$. Then

$$\begin{aligned}\langle a,b\rangle_K\langle a',b\rangle_K &= (-1)^{mn(q-1)/2}\chi(\bar{\varepsilon})^n\chi(\bar{\delta})^m(-1)^{m'n(q-1)/2}\chi(\bar{\varepsilon'})^n\chi(\bar{\delta})^{m'}\\ &= (-1)^{(m+m')n(q-1)/2}(\chi(\bar{\varepsilon})\chi(\bar{\varepsilon'}))^n\chi(\bar{\delta})^{m+m'}\\ &= (-1)^{(m+m')n(q-1)/2}\chi(\bar{\varepsilon}\bar{\varepsilon'})^n\chi(\bar{\delta})^{m+m'}\\ &= \langle aa',b\rangle_K\end{aligned}$$

so the formula is multiplicative. We now show $(a,b)_K=\langle a,b\rangle_K$ for all $a,b\in K^\times$. By the algebraic properties just proved, it is enough to check equality for a,b equal to π or units. So we have three cases to check.

Case 1: (a,b are units) Let $a=\varepsilon$ and $b=\delta$. Then $(\varepsilon,\delta)_K=1$ by Theorem 3.37. On the other hand,

$$\langle\varepsilon,\delta\rangle_K=(-1)^0\chi(\bar{\varepsilon})^0\chi(\bar{\delta})^0=1.$$

So the formula is valid for two units.

Case 2: ($a=\pi$ and $b=\varepsilon$ is a unit) We have $(\pi,\varepsilon)_K=1$ if and only if the equation $\pi x^2+\varepsilon y^2-z^2=0$ has a non-trivial solution. But, by Theorem 2.38 and Theorem 3.34, this has a non-trivial solution if and only if ε is a square in K. This last statement is equivalent to ε being a square in the residue field. Similarly,

$$\langle\pi,\varepsilon\rangle_K=(-1)^0\chi(1)^0\chi(\bar{\varepsilon})=\chi(\bar{\varepsilon}),$$

which is 1 if and only if ε is a square in the residue field. So the formula holds in this case too.

Case 3: ($a=b=\pi$) In this case, $(\pi,\pi)_K=(\pi,-1)_K=\langle\pi,-1\rangle_K$ by Example 3.52 and Case 2. Moreover,

$$\langle\pi,-1\rangle_K=\chi(-1)=(-1)^{(q-1)/2}=\langle\pi,\pi\rangle_K$$

so $(\pi,\pi)_K=\langle\pi,\pi\rangle_K$. □

Theorem 3.54. *For $a,b\in\mathbb{R}^\times$,*

$$(a,b)_{\mathbb{R}}=\begin{cases}1, & \text{if } a>0 \text{ or } b>0,\\ -1, & \text{if } a<0 \text{ and } b<0.\end{cases}$$

Proof. This follows by a direct calculation. □

For an explicit formula for $(a, b)_K$ when $K = \mathbb{Q}_2$ see [8, p. 20]. We do not need such a formula in our work.

We now discuss some applications of the Hilbert symbol. In Chapter 1 we met two invariants for quadratic forms, the dimension and the discriminant. These were enough to classify non-degenerate quadratic forms over $\mathbb{C}$ and $\mathbb{F}$. The classification of non-degenerate quadratic forms over $\mathbb{R}$ needed an additional invariant (the number of positive coefficients in a diagonalization). Here we introduce an invariant for quadratic forms over local fields using the Hilbert symbol.

Definition 3.55. Let K be a local field not of characteristic 2 or $K = \mathbb{R}$. Let Q be a non-degenerate, n-dimensional quadratic form over K that is equivalent to the diagonal form

$$a_1 x_1^2 + \cdots + a_n x_n^2.$$

Then

$$c_K(Q) = \prod_{i<j} (a_i, a_j)_K \in \{\pm 1\}$$

is called the *Hasse invariant* of Q. If Q is 1-dimensional then we make the convention that $c_K(Q) = 1$.

Remark 3.56. Notice that the Hasse invariant is only defined for non-degenerate quadratic forms since the Hilbert symbol $(a, b)_K$ is undefined if a or b is 0.

Before using the Hasse invariant, we show that it only depends on the equivalence class of a quadratic form and not on a particular diagonalization (of course it would be silly to name it an invariant if this was not true). The following two lemmas are needed for this task.

Lemma 3.57. *Let Q be a non-degenerate quadratic form over a vector space V. Let $E = \{e_1, \ldots, e_n\}$ and $E' = \{e'_1, \ldots, e'_n\}$ be any pair of orthogonal bases of V. Then there is a sequence of orthogonal bases*

$$E_i = \{e_{i1}, \ldots, e_{in}\},$$

for $i = 1, \ldots, m$ such that:

1) $E_1 = E$ *and* $E_m = E'$,

2) *for each* i, *the bases* E_i, E_{i+1} *have at least* $n-2$ *elements in common* (*i.e.,* $e_{ij} = e_{(i+1)j}$ *for at least* $n-2$ *values of* j).

Proof. Let F be the field over which V is a vector space. If E' is a permutation of E then Lemma 3.57 is trivially true since all permutations are products of transpositions. Thus, we do not have to take the order of a basis into account.

We now proceed by induction on the dimension of V. When $\dim V = 1$ or 2 the lemma trivially holds. For $n \geq 3$, assume that it is true when $\dim V < n$. We want to show that there is a sequence of the desired type from E to an orthogonal basis $\{e'_1, e^*_2, \ldots, e^*_n\}$ for some $e^*_i \in V$. To do so, let $\{f_1, \ldots, f_n\}$ be an orthogonal basis such that there is a sequence from E to it fitting the second condition of the lemma and that e'_1 is a linear combination of the minimum number of basis vectors possible. We will show that this minimum is 1.

Assume without loss of generality e'_1 is a linear combination of the first J basis vectors, so explicitly

$$e'_1 = \sum_{i=1}^{J} a_i f_i$$

where each $a_i \neq 0$.

Suppose $J > 1$. We first show that there are always two of these vectors such that $Q(a_j f_j + a_k f_k) \neq 0$. Assume, without loss of generality, that

$$\begin{aligned} 0 = Q(a_1 f_1 + a_2 f_2) &= a_1^2 Q(f_1) + a_2^2 Q(f_2), \\ 0 = Q(a_1 f_1 + a_3 f_3) &= a_1^2 Q(f_1) + a_3^2 Q(f_3), \\ 0 = Q(a_2 f_2 + a_3 f_3) &= a_2^2 Q(f_2) + a_3^2 Q(f_3). \end{aligned}$$

Then, $Q(f_1) = Q(f_2) = Q(f_3) = 0$. This is a contradiction since $Q(f_i) \neq 0$ (V is non-degenerate). Therefore, there is always a pair of vectors f_j, f_k with $1 \leq j < k \leq J$ such that $Q(a_j f_j + a_k f_k) \neq 0$. By permuting the basis vectors, we may assume $Q(a_1 f_1 + a_2 f_2) \neq 0$.

We now want to find a non-zero vector orthogonal to $a_1 f_1 + a_2 f_2$ and $f_3, \ldots, f_n$. For unknowns b_1 and b_2 in F, suppose

$$\begin{aligned} 0 &= \mathrm{B}(a_1 f_1 + a_2 f_2, b_1 f_1 + b_2 f_2) \\ &= a_1 b_1 Q(f_1) + a_2 b_2 Q(f_2) \end{aligned}$$

where B is the bilinear form associated to Q. Taking $b_1 = a_2 Q(f_2)$ and $b_2 = -a_1 Q(f_1)$ gives the desired non-zero vector. Let

$$\begin{aligned} \tilde{f}_1 &= a_1 f_1 + a_2 f_2 \\ \tilde{f}_2 &= a_2 Q(f_2) f_1 - a_1 Q(f_1) f_2 \\ \tilde{f}_j &= f_j \text{ for } j = 3, \ldots, n. \end{aligned}$$

Then $\mathrm{B}(\tilde{f}_i, \tilde{f}_j) = 0$ for all $i \neq j$ in $\{1, \ldots, n\}$ and note that

$$\begin{aligned} Q(\tilde{f}_2) &= Q(a_2 Q(f_2) f_1 - a_1 Q(f_1) f_2) \\ &= Q(f_1) Q(f_2) (a_1^2 Q(f_1) + a_2^2 Q(f_2)) \\ &= Q(f_1) Q(f_2) Q(a_1 f_1 + a_2 f_2) \\ &\neq 0 \end{aligned}$$

so $\{\tilde{f}_1, \ldots, \tilde{f}_n\}$ is an orthogonal basis. Furthermore $e_1' = \tilde{f}_1 + \sum_{i=3}^{J} a_i \tilde{f}_i$ (the sum is empty if $J = 2$), a total of $J-1$ vectors, contradicting the assumption that J is minimal. Hence $J = 1$. Thus there is a sequence of the desired type starting with E and ending with an orthogonal basis $\{e_1', e_2^*, \ldots, e_n^*\}$ for some $e_i^* \in V$.

The inductive step comes now. Since Fe_1' is a non-degenerate subspace of V, $e_1'^{\perp}$ is also non-degenerate (because V is non-degenerate). Moreover, $e_1'^{\perp}$ is an $(n-1)$-dimensional subspace of V spanned by $\{e_2^*, \ldots, e_n^*\}$. By the induction hypothesis, there is a sequence of orthogonal bases of $e_1'^{\perp}$ starting with $\{e_2^*, \ldots, e_n^*\}$ and ending with $\{e_2', \ldots, e_n'\}$ fitting the necessary conditions. Attaching e_1' to this sequence of bases gives a sequence of the desired type from $\{e_1', e_2^*, \ldots, e_n^*\}$ to E' and thus from E to E'. □

Lemma 3.58. *Let K be a local field not of characteristic 2 or $K = \mathbb{R}$. Let Q be a non-degenerate, 2-dimensional quadratic form over K. Then for $b \in K^\times$, Q has the value b over K if and only if $(b, -\operatorname{disc}(Q))_K = c_K(Q)$.*

Proof. Without loss of generality assume Q is diagonal:

$$Q(x_1, x_2) = a_1 x_1^2 + a_2 x_2^2$$

for $a_i \in K^\times$. Then $Q(x_1, x_2) = b$ is solvable over K if and only if the quadratic form $Q(x_1, x_2) - b x_3^2$ has a null vector. This last statement is equivalent to the condition $(a_1/b, a_2/b)_K = 1$.

Using the algebraic properties of the Hilbert symbol (Corollary 3.50 and Example 3.52),

$$\begin{aligned}(a_1/b, a_2/b)_K &= (b, a_2)_K(b, a_1)_K(b, b)_K(a_1, a_2)_K \\ &= (b, -a_1a_2)_K(a_1, a_2)_K \\ &= (b, -\operatorname{disc}(Q))_K(a_1, a_2)_K.\end{aligned}$$

So $(a_1/b, a_2/b)_K = 1$ if and only if $(b, -\operatorname{disc}(Q))_K = (a_1, a_2)_K$. □

Remark 3.59. Lemma 3.58 tells us that $c_K(Q)$ is independent of a choice of diagonalization when $\dim Q = 2$ since $(b, -\operatorname{disc}(Q))_K$ does not depend on a particular diagonalization.

Theorem 3.60. *Let K be a local field not of characteristic 2 or $K = \mathbb{R}$. If*

$$Q(x_1, \ldots, x_n) = a_1x_1^2 + \cdots + a_nx_n^2$$

is a non-degenerate diagonalized quadratic form that is equivalent to

$$Q'(x_1, \ldots, x_n) = b_1x_1^2 + \cdots + b_nx_n^2$$

then $c_K(Q) = c_K(Q')$.

Proof. The theorem is vacuously true for $n = 1$. For $n = 2$ it is true by Lemma 3.58. So suppose $n > 2$. By Lemma 3.57 there is a sequence of orthogonal bases that takes $Q(x_1, \ldots, x_n)$ to $Q'(x_1, \ldots, x_n)$ such that each basis in the sequence differs from the previous by at most two vectors. If we can show that $c_K(Q)$ is invariant after one step in such a sequence, then it will be invariant throughout all steps in the sequence. Since the coefficients in a diagonalization of a quadratic form are its values on an orthogonal basis, it is sufficient to assume that a_i does not equal some b_j for at most two values of i. What is more, $\prod_{i<j}(a_i, a_j)_K$ is independent of any permutation of $a_1, \ldots, a_n$. So without loss of generality we may assume $a_i = b_i$ for $i > 2$. Then $a_1x_1^2 + a_2x_2^2$ is equivalent to $b_1x_1^2 + b_2x_2^2$ by Witt cancellation, so $a_1a_2 \sim b_1b_2$ and

$$(a_1, a_2)_K = (b_1, b_2)_K$$

from the $n = 2$ case of this theorem. Now compute the Hasse invariants of

Q and Q':

$$\begin{aligned} c_K(Q) &= (a_1,a_2)_K \prod_{2<i}(a_1a_2,a_i)_K \prod_{2<i<j}(a_i,a_j)_K \\ &= (b_1,b_2)_K \prod_{2<i}(b_1b_2,b_i)_K \prod_{2<i<j}(b_i,b_j)_K \\ &= c_K(Q'). \end{aligned}$$

Thus the Hasse invariant is independent of a particular choice of diagonalization. □

Lemma 3.58 already illustrates how the Hasse invariant plays an important role in finding necessary and sufficient conditions for when a 2-dimensional quadratic form over a local field represents a particular value in the field. As it turns out, the Hasse invariant also plays a major role in the following 3-dimensional analogue of Lemma 3.58.

Theorem 3.61. *Let K be a local field not of characteristic 2 or $K = \mathbb{R}$. A non-degenerate 3-dimensional quadratic form Q over K takes on a value $b \in K^\times$ if and only if the 4-dimensional quadratic form $Q - bx^2$ satisfies at least one of the following conditions:*

1) $\operatorname{disc}(Q - bx^2) \notin K^{\times^2}$,

2) $c_K(Q - bx^2) = (-1,-1)_K$.

Proof. Without loss of generality assume that Q is diagonalized. Write $b = a_4$ and $Q = a_1x_1^2 + a_2x_2^2 - a_3x_3^2$, so

$$Q(x_1,x_2,x_3) - a_4x_4^2 = Q_1(x_1,x_2) - Q_2(x_3,x_4)$$

where

$$Q_1(x_1,x_2) = a_1x_1^2 + a_2x_2^2$$

and

$$Q_2(x_3,x_4) = a_3x_3^2 + a_4x_4^2.$$

For simplicity, write $(a,b)_K$ as (a,b).

Case 1: ($-\operatorname{disc}(Q_1)$ or $-\operatorname{disc}(Q_2)$ is in $K^{\times^2}$) Without loss of generality we may assume that $-\operatorname{disc}(Q_1)$ is a square in $K^\times$. Then Q_1 has a null vector so $Q - a_4x_4^2$ has a null vector too. Thus Q has the value $a_4 = b$. We now

show that $Q - a_4x_4^2$ satisfies at least one of (1) and (2). Since $-\operatorname{disc}(Q_1)$ is in $K^{\times^2}$ we have

$$\operatorname{disc}(Q - a_4x_4^2) = \operatorname{disc}(Q_1)\operatorname{disc}(Q_2) \sim -\operatorname{disc}(Q_2).$$

Suppose $-\operatorname{disc}(Q_2)$ is a square in $K^\times$ too. Otherwise (1) is satisfied. We will show that $c_K(Q - a_4x_4^2)$ is $(-1,-1) = (-1,-1)_K$.

By a direct calculation,

$$\begin{aligned} c_K(Q - a_4x_4^2) &= (a_1,a_2)(a_1,-a_3)(a_1,-a_4)(a_2,-a_3)(a_2,-a_4)(-a_3,-a_4) \\ &= (a_1,a_2)(a_1,a_3a_4)(a_2,a_3a_4)(-a_3,-a_4) \\ &= (a_1,a_2)(a_1a_2,a_3a_4)(-a_3,-a_4). \end{aligned}$$

When $-\operatorname{disc}(Q_1) = -a_1a_2$ and $-\operatorname{disc}(Q_2) = -a_3a_4$ are in $K^{\times^2}$,

$$\begin{aligned} c_K(Q - a_4x_4^2) &= (a_1,a_2)(a_1a_2,a_3a_4)(-a_3,-a_4) \\ &= (a_1,-a_1)(-1,-1)(-a_3,a_3) \\ &= (-1,-1) \end{aligned}$$

just as we wanted.

Case 2: ($-\operatorname{disc}(Q_1)$ and $-\operatorname{disc}(Q_2)$ are not squares in $K^\times$) We show this case by proving the equivalence of the negations. That is, $Q - bx^2$ does not have a null vector if and only if $Q - bx^2$ satisfies neither (1) nor (2).

By Corollary 2.45, $Q - a_4x_4^2$ has a null vector if and only if Q_1 and Q_2 take on a common non-zero value. From Lemma 3.58 we have that Q_1 and Q_2 represent a common non-zero value, say $d \in K^\times$, if and only if

$$(d, -\operatorname{disc}(Q_1)) = (a_1,a_2) \tag{3.13}$$

and

$$(d, -\operatorname{disc}(Q_2)) = (a_3,a_4). \tag{3.14}$$

Since $-\operatorname{disc}(Q_1)$ and $-\operatorname{disc}(Q_2)$ are not in $K^{\times^2}$, $K(\sqrt{-a_1a_2})$ and $K(\sqrt{-a_3a_4})$ are quadratic extensions of K with $x^2 + \operatorname{disc}(Q_1)y^2$ and $x^2 + \operatorname{disc}(Q_2)y^2$ as their respective norm forms. By the definition and symmetry of the Hilbert symbol,

$$(d, -\operatorname{disc}(Q_i)) = 1$$

if and only if the norm form $x^2 + \operatorname{disc}(Q_i)y^2$ represents d for some $x, y \in K$. By Theorem 3.48 the subgroup of values represented by each of these norm

forms has index two in $K^\times$. So the set of d satisfying each of (3.13) and (3.14) lies in exactly half of the cosets of $K^{\times^2}$. Hence $Q - a_4x_4^2$ not having a null vector is equivalent to the sets of d satisfying (3.13) and (3.14) being complementary. If

$$\operatorname{disc}(Q_1) \sim \operatorname{disc}(Q_2) \tag{3.15}$$

and

$$(a_1, a_2) = -(a_3, a_4) \tag{3.16}$$

then it is clear that the sets of d satisfying (3.13) and (3.14) are complementary. We now prove the converse.

Suppose that $(a_1, a_2) = (a_3, a_4)$. Then when $d = 1$

$$(d, -\operatorname{disc}(Q_1)) = (d, -\operatorname{disc}(Q_2)).$$

Thus for $d = 1$, either (3.13) and (3.14) are true or both of them are false. Hence $(a_1, a_2) = -(a_3, a_4)$ if the sets of d satisfying (3.13) and (3.14) are complementary. If $\operatorname{disc}(Q_1) \not\sim \operatorname{disc}(Q_2)$ then $\operatorname{disc}(Q_1)/\operatorname{disc}(Q_2)$ is not a square in $K^\times$. This means there is some d_0 such that $(d_0, \operatorname{disc}(Q_1)/\operatorname{disc}(Q_2)) = -1$ by (2) of Corollary 3.50. Thus $(d_0, -\operatorname{disc}(Q_1)) = -(d_0, -\operatorname{disc}(Q_2))$ while $(a_1, a_2) = -(a_3, a_4)$. So for $d = d_0$ either (3.13) and (3.14) are true or both of them are false. Hence the sets of d satisfying (3.13) and (3.14) being complementary is equivalent to (3.15) and (3.16) being true.

So we now have that $Q - a_4x_4^2$ does not have a null vector if and only if (3.15) and (3.16) are satisfied. We will show that (1) and (2) fail if and only if (3.15) and (3.16) hold. That is, $a_1a_2a_3a_4 \sim 1$ and $c_K(Q - a_4x_4^2) = -(-1, -1)$ if and only if (3.15) and (3.16) are true. It is clear that $a_1a_2a_3a_4 \sim 1$ is equivalent to (3.15) being true. So let us assume that $a_1a_2a_3a_4 \sim 1$ and show that $c_K(Q - a_4x_4^2) = -(-1, -1)$ if and only if $(a_1, a_2) = -(a_3, a_4)$. From before,

$$c_K(Q - a_4x_4^2) = (a_1, a_2)(a_1a_2, a_3a_4)(-a_3, -a_4).$$

Since $a_1a_2a_3a_4 \sim 1$,

$$\begin{aligned}
(a_1, a_2)(a_1a_2, a_3a_4)(-a_3, -a_4) &= (a_1, a_2)(a_3a_4, a_3a_4)(-a_3, -a_4) \\
&= (a_1, a_2)(-1, a_3a_4)(-a_3, -a_4) \\
&= (a_1, a_2)(-1, a_3)^2(-1, a_4)^2(-1, -1)(a_3, a_4) \\
&= (a_1, a_2)(-1, -1)(a_3, a_4),
\end{aligned}$$

which is $-(-1, -1)$ if and only if $(a_1, a_2) = -(a_3, a_4)$. □

Corollary 3.62. *Let K be a local field not of characteristic 2 or $K = \mathbb{R}$. A non-degenerate, 3-dimensional quadratic form over K takes on every value in all but possibly one coset of $K^\times/K^{\times^2}$.*

Proof. Let Q be a non-degenerate, 3-dimensional quadratic form over K. If Q does not take on some value, say b, then $Q - bx^2$ does not have a null vector and hence $\operatorname{disc}(Q - bx^2) = -b\operatorname{disc}(Q)$ is a square in $K^\times$. So the only values that Q might not take on are in the same square class as $-\operatorname{disc}(Q)$. For $K = \mathbb{R}$, this corollary is trivial since $\mathbb{R}^\times/\mathbb{R}^{\times^2}$ has size 2. □

We can now extend Corollary 3.39 to 2-adic fields. Our argument will give a uniform treatment for all local fields not of characteristic 2.

Theorem 3.63. *Let K be a local field not of characteristic 2. Every quadratic form of dimension ≥ 5 over K has a null vector.*

Proof. Let Q be a quadratic form over K of dimension $n \geq 5$. We may assume that Q is non-degenerate since any degenerate quadratic form has a null vector (Theorem 2.32). Furthermore, we can assume Q is diagonalized. For $n = 5$, write

$$Q(x_1, x_2, x_3, x_4, x_5) = Q_1(x_1, x_2, x_3) - Q_2(x_4, x_5)$$

where

$$Q_1(x_1, x_2, x_3) = a_1x_1^2 + a_2x_2^2 + a_3x_3^2$$

and

$$Q_2(x_4, x_5) = -a_4x_4^2 - a_5x_5^2$$

for $a_i \in K^\times$. Then it is equivalent by Corollary 2.45 to show that Q_1 and Q_2 take on a common non-zero value. By Corollary 3.62, Q_1 takes on all values of $K^\times$ except in possibly one coset of $K^\times/K^{\times^2}$. If $-\operatorname{disc}(Q_2)$ is a square in $K^\times$ then Q_2 is universal by Lemma 3.58. On the other hand, if $-\operatorname{disc}(Q_2)$ is not a square in $K^\times$ then $-a_4Q_2$ is equivalent to a norm form of a quadratic extension of K. By Theorem 3.48 the set of non-zero values of $-a_4Q_2$ is a subgroup of index 2 in $K^\times$. Hence no matter what the value of $-\operatorname{disc}(Q_2)$ is, Q_2 takes on all values in at least half of the cosets of $K^\times/K^{\times^2}$. In any local field, $[K^\times : K^{\times^2}] > 2$ by Theorem 3.29 (this is false for $K = \mathbb{R}$) so Q_1 and Q_2 must take on a common non-zero value.

For $n > 5$, set all but five of the variables equal to zero, reducing to the $n = 5$ case. □

Remark 3.64. One nice result of Theorem 3.63 is that every non-degenerate quadratic form over a local field of dimension greater than or equal to four is universal. This is false for quadratic forms over $\mathbb{R}$; a sum of squares is not universal over $\mathbb{R}$.

The last theorem in this chapter characterizes non-degenerate quadratic forms over local fields up to equivalence using the Hasse invariant together with the dimension and discriminant.

Theorem 3.65. *The dimension, discriminant, and Hasse invariant determine a non-degenerate quadratic form over a local field up to equivalence.*

Proof. Let Q_1 and Q_2 be n-dimensional non-degenerate quadratic forms over K such that $\text{disc}(Q_1) \sim \text{disc}(Q_2)$ and $c_K(Q_1) = c_K(Q_2)$. We show that $Q_1 \cong Q_2$ by inducting on the dimension. When $n = 1$ the theorem is trivially true. Now suppose $n > 1$. We want to show that Q_1 and Q_2 take on a common value. Once this is shown, we will diagonalize and use induction.

For $n = 2$, Q_1 and Q_2 take the same values by Lemma 3.58. For $n > 2$, the quadratic form $Q_1 \perp -Q_2$ has dimension $2n > 5$ so by Corollary 2.45 and Theorem 3.63, Q_1 and Q_2 take on a common non-zero value. Let this common value be d. Then diagonalizing gives

$$Q_1(x_1, \ldots, x_n) = dx_1^2 + a_2x_2^2 + \cdots + a_nx_n^2$$

and (in a possibly different set of coordinates)

$$Q_2(y_1, \ldots, y_n) = dy_1^2 + b_2y_2^2 + \cdots + b_ny_n^2$$

where $a_i, b_i \in K^\times$. Denote $a_2x_2^2 + \cdots + a_nx_n^2$ by Q_1' and $b_2y_2^2 + \cdots + b_ny_n^2$ by Q_2'. It is clear that Q_1' and Q_2' have dimension $n - 1$. Furthermore, $\text{disc}(Q_1') \sim \text{disc}(Q_2')$ since

$$da_2 \cdots a_n \sim db_2 \cdots b_n$$

(just cancel the d's). Lastly we show that $c_K(Q_1') = c_K(Q_2')$. The Hasse invariant of Q_1 is

$$\begin{aligned} c_K(Q_1) &= \prod_i (d, a_i)_K \prod_{i<j} (a_i, a_j)_K \\ &= (d, \text{disc}(Q_1'))_K c_K(Q_1') \end{aligned}$$

and the Hasse invariant of Q_2 is

$$\begin{aligned} c_K(Q_2) &= \prod_i (d, b_i)_K \prod_{i<j} (b_i, b_j)_K \\ &= (d, \operatorname{disc}(Q_2'))_K c_K(Q_2'). \end{aligned}$$

So $c_K(Q_1') = c_K(Q_2')$ since $c_K(Q_1) = c_K(Q_2)$ and $\operatorname{disc}(Q_1') \sim \operatorname{disc}(Q_2')$. Then by induction $Q_1' \cong Q_2'$ and hence $Q_1 \cong Q_2$. □

Chapter 4

Hasse–Minkowski Over $\mathbb{Q}$

We are now ready to prove the Hasse–Minkowski theorem over $\mathbb{Q}$. After doing so we will discuss some applications. We restate the Hasse–Minkowski theorem:

Theorem 4.1. *Let $Q(x_1, \ldots, x_n)$ be a non-degenerate quadratic form over $\mathbb{Q}$. The equation $Q(x_1, \ldots, x_n) = 0$ is non-trivially solvable over $\mathbb{Q}$ if and only if it is non-trivially solvable over $\mathbb{R}$ and every $\mathbb{Q}_p$.*

Theorems 1.16 and 4.1 are equivalent. Clearly Theorem 4.1 is a special case of Theorem 1.16. Suppose, conversely, that Theorem 4.1 is true. Let Q be a quadratic form over $\mathbb{Q}$. If Q is non-degenerate then the second part of Theorem 1.16 is true for Q since this is Theorem 4.1. The first part of Theorem 1.16 for Q follows from Theorem 4.1 for $Q - rx_{n+1}^2$ by Corollary 2.43.

If Q is degenerate then part 2 of Theorem 1.16 is true by Theorem 2.32. We show that part 1 of Theorem 1.16 follows from Theorem 4.1. It is sufficient to prove that $Q = r$ over $\mathbb{Q}$ if $Q = r$ over $\mathbb{R}$ and all $\mathbb{Q}_p$ (of course granting that Theorem 4.1 is true) since the other direction is obvious. Let V be the vector space on which Q is defined. So $V^\perp \neq \{0\}$ because Q is degenerate. Write (in a non-canonical manner) $V = W \perp V^\perp$ and $Q' = Q|_W$. Fix a basis of V that respects this orthogonal decomposition of V. Thus Q' is a non-degenerate quadratic form over $\mathbb{Q}$ and $Q(w, u) = Q'(w)$ where $w \in W$ and $u \in V^\perp$. The equation $Q'(w) - rt^2 = 0$ is non-trivially solvable over $\mathbb{R}$ and all $\mathbb{Q}_p$ by Corollary 2.43 since $Q(w, u) = r$ is solvable over $\mathbb{R}$ and all $\mathbb{Q}_p$. Theorem 4.1 implies that $Q'(w) - rt^2 = 0$ has a non-trivial solution over $\mathbb{Q}$ so $Q(w, 0) = Q'(w) = r$ has a solution over $\mathbb{Q}$ by Corollary 2.43.

The proof of Hasse–Minkowski uses induction on the dimension n of the quadratic form. The case $n = 1$ is trivial so we begin with $n = 2$. Recall that the collection of completions of $\mathbb{Q}$ are denoted $\mathbb{Q}_v$ with v equal to a prime p in $\mathbb{Z}$ or $v = \infty$ ($\mathbb{Q}_\infty = \mathbb{R}$).

We can restrict our attention to diagonalized quadratic forms since Theorem 2.22 tells us that any quadratic form can be diagonalized.

4.1 $n = 2$

The statement of Hasse–Minkowski in dimension 2 says: for $a, b \in \mathbb{Q}^\times$, the equation $ax^2 + by^2 = 0$ is non-trivially solvable over $\mathbb{Q}$ if and only if it is non-trivially solvable over all $\mathbb{Q}_v$.

Lemma 4.2. *Over any field F (not of characteristic 2) the non-degenerate quadratic form $ax^2 + by^2$ has a null vector over F if and only if $-b/a \in F^{\times 2}$.*

Proof. Suppose the equation $ax^2 + by^2 = 0$ is non-trivially solvable over F. Let (x_0, y_0) be a non-trivial solution. If $x_0 = 0$ then $y_0 = 0$ since $by_0^2 = 0$ and $b \neq 0$. Similarly, if $y_0 = 0$ then $x_0 = 0$. So both x_0 and y_0 are non-zero. Thus,

$$\frac{x_0^2}{y_0^2} = -\frac{b}{a}$$

so $-b/a$ is in $F^{\times 2}$.

Conversely, if $-b/a = c^2$ for some $c \in F^\times$ then $(c, 1)$ is a null vector. □

So our proof is reduced to the next result, called the square theorem.

Theorem 4.3. *Let $c \in \mathbb{Q}^\times$. Then c is a square in $\mathbb{Q}$ if and only if c is a square in all $\mathbb{Q}_v$.*

Proof. Clearly only one direction needs proving since $\mathbb{Q}$ is a subset of each $\mathbb{Q}_v$. Write $c = \pm p_1^{e_1} \dots p_r^{e_r}$ for distinct primes p_i. Suppose c is a square in all $\mathbb{Q}_v$. Then $\mathrm{ord}_{p_i}(c)$ is even for every p_i so e_i is even for all i and $c > 0$ since it is a square in $\mathbb{R}$. Hence c is a square in $\mathbb{Q}$. □

4.2 $n = 3$

We want to show for a, b, c in $\mathbb{Q}^\times$ that if $ax^2 + by^2 + cz^2 = 0$ non-trivially over all $\mathbb{Q}_v$ then $ax^2 + by^2 + cz^2 = 0$ non-trivially over $\mathbb{Q}$. The converse is trivial.

First we make a series of reductions. The coefficients a, b, and c may be taken to be integers since scaling by a common denominator of the coefficients does not affect the existence of a non-trivial rational solution. If any of the coefficients have a square factor then by a linear change of variables we can get a, b and c to be square-free. For instance, if a square integer m^2 is a factor of a then let $x' = mx$ and $a' = a/m^2$, so

$$ax^2 + by^2 + cz^2 = a'(mx)^2 + by^2 + cz^2.$$

Since $ax^2 + by^2 + cz^2 = 0$ has a non-trivial solution in $\mathbb{R}$, one of the coefficients must have a different sign from the other two. So without loss of generality assume $a, b > 0$ and $c < 0$. Rewrite c as $-c$ with $c > 0$, so our quadratic form becomes $ax^2 + by^2 - cz^2$. Lastly, we can multiply by c and then use a linear change of variables to get the coefficients of x and y to be square-free integers again while the coefficient of z is -1. Thus we can assume our quadratic form is $ax^2 + by^2 - z^2$ where a and b are square-free integers. We now give a proof by induction on the size of the coefficients.

Theorem 4.4. *Let a and b be square-free integers. If $ax^2 + by^2 = z^2$ has a non-trivial solution in every $\mathbb{Q}_v$, then it has a non-trivial rational solution.*

Proof. We will induct on $|a| + |b|$. Let $|a| + |b| = 2$. Then $|a| = |b| = 1$. So the equations that fit these conditions are the following:

- $x^2 + y^2 - z^2 = 0$ which has the solution $(1, 0, 1)$,
- $x^2 - y^2 - z^2 = 0$ which has the solution $(1, 0, 1)$,
- $-x^2 + y^2 - z^2 = 0$ which has the solution $(1, 1, 0)$,
- $-x^2 - y^2 - z^2 = 0$ which only has the trivial solution in $\mathbb{R}$ and $\mathbb{Q}_2$ as shown in Example 3.43. (By Example 3.38 there are non-trivial solutions in $\mathbb{Q}_p$ for all odd p.)

Thus the theorem holds for $|a| + |b| = 2$. Now assume that $|a| + |b| > 2$ and that the theorem is true for all smaller values. Without loss of generality, $|a| \leq |b|$, so $|b| \geq 2$.

First we will show that a is a square $\bmod\ b$. Since b is square-free, by the Chinese remainder theorem it is sufficient to show that a is a square mod p for each prime p dividing b.

By our assumption, there is a non-trivial solution $(x_p, y_p, z_p) \in \mathbb{Q}_p^3$ for each p. So $ax_p^2 + by_p^2 = z_p^2$. Then by scaling we can assume that (x_p, y_p, z_p) is a primitive solution (meaning that $x_p, y_p, z_p \in \mathbb{Z}_p$ and at least one of x_p, y_p, z_p is in $\mathbb{Z}_p^\times$). We will show that x_p is in $\mathbb{Z}_p^\times$. If $p|x_p$, then $0 \equiv z_p^2 \bmod p$ since $p|b$. Hence $p|z_p$. This means $by_p^2 \equiv 0 \bmod p^2$. So $p|y_p$ because b is square-free. Thus $p|x_p$, $p|y_p$ and $p|z_p$ which contradicts the assumption that (x_p, y_p, z_p) is primitive. Since p divides b, we have $ax_p^2 \equiv z_p^2 \bmod p$, and $x_p \in \mathbb{Z}_p^\times$, so a is a square modulo p.

Write $a = m^2 \bmod b$. Without loss of generality, $|m| \leq \frac{1}{2}|b|$. Then $a + bq = m^2$ for some $q \in \mathbb{Z}$. We assumed a to be square-free, and the case $a = 1$ is trivial, so $q \neq 0$. Write $q = cd^2$ where c is a square-free integer. Dividing $a + bcd^2 = m^2$ by $(cd)^2$ gives

$$\begin{aligned} \frac{b}{c} &= \frac{m^2 - a}{(cd)^2} \\ &= \frac{m^2}{(cd)^2} - a\frac{1}{(cd)^2}, \end{aligned} \tag{4.1}$$

so b/c is in $\{x^2 - ay^2 \neq 0 : x, y \in \mathbb{Q}\}$. By the hypothesis of the theorem and Theorem 3.47, $b = x_v^2 - ay_v^2$ in each $\mathbb{Q}_v$, so (4.1) shows that c has this same form in each $\mathbb{Q}_v$ because $\{x^2 - ay^2 \neq 0 : x, y \in \mathbb{Q}_v\}$ is a subgroup of $\mathbb{Q}_v^\times$. Hence there are non-trivial solutions to $ax^2 + cy^2 = z^2$ in all $\mathbb{Q}_v$ (Theorem 3.47 again).

We now show $|a| + |c| < |a| + |b|$ or in other words $|c| < |b|$. Indeed, since $|a| \leq |b|$,

$$\begin{aligned} |bcd^2| &= |m^2 - a| \\ &\leq |m|^2 + |a| \\ &\leq \frac{1}{4}|b|^2 + |b|. \end{aligned}$$

So $|c| \leq |cd^2| \leq \frac{1}{4}|b| + 1$. This means $|c| < |b|$ since $|b| \geq 2$. Then by the induction hypothesis there is a non-trivial rational solution to $ax^2 + cy^2 = z^2$. Thus $c = r^2 - as^2$ for some $r, s \in \mathbb{Q}$. Again by (4.1), b has the form $x^2 - ay^2$ for some $x, y \in \mathbb{Q}$. So $ax^2 + by^2 = z^2$ has a non-trivial rational solution. $\square$

For a different proof see [1, pp. 62 – 64] or [2, pp. 78 – 82]. The proof above is based on [8, p. 42] and will easily generalize to $\mathbb{F}(T)$ later.

Both the $n = 2$ and $n = 3$ cases for Hasse–Minkowski over $\mathbb{Q}$ have been relatively non-technical. This will change in the $n = 4$ case. In contrast, the $n = 2$ and $n = 3$ cases for the Hasse–Minkowski theorem over all number fields are very hard, whereas the $n = 4$ case is a purely algebraic corollary of the $n = 3$ case. Let us see what this means.

Every finite extension of $\mathbb{Q}$ has its own set of archimedean and non-archimedean absolute values that lead to completions that are finite extensions of $\mathbb{Q}_p$ and $\mathbb{R}$. The following theorem gives the statement for Hasse–Minkowski over any number field.

Theorem 4.5. *Let K be any finite extension of $\mathbb{Q}$. Let Q be a quadratic form over K of any dimension $n \geq 1$. Then*

1) *for $r \in K^\times$, $Q(x_1, \ldots, x_n) = r$ is solvable over K if and only if it is solvable over each K_v,*

2) *the equation $Q(x_1, \ldots, x_n) = 0$ is non-trivially solvable over K if and only if it is non-trivially solvable over each K_v.*

Remark 4.6. Theorem 4.5 can be reduced to the case of null vectors for non-degenerate quadratic forms in the same manner as over $\mathbb{Q}$.

We show the algebraic step of using the $n = 3$ case of Theorem 4.5 to prove the $n = 4$ case of Theorem 4.5.

Theorem 4.7. *If the Hasse–Minkowski theorem is true in dimension 3 over all number fields then it is true in dimension 4 over all number fields.*

Proof. Let K be any number field and Q be a 4-dimensional quadratic form over K with a null vector over each K_v. We want to show that Q has a null vector over K. Without loss of generality, Q is non-degenerate. So we may assume $Q = ax^2 + by^2 + cz^2 + dt^2$ where a, b, c, d are in $K^\times$.

Case 1: Let $\operatorname{disc}(Q)$ be a square in $K^\times$. Then scaling Q by a gives $\operatorname{disc}(aQ) = a^4 \operatorname{disc}(Q)$ so $\operatorname{disc}(aQ)$ is still a square. Furthermore, since aQ is equivalent to $x^2 + aby^2 + acz^2 + adt^2$, assume without loss of generality that $a = 1$. Thus $\operatorname{disc}(Q) = bcd$ is a square in $K^\times$. This means $d \sim bc$. So Q is equivalent to

$$x^2 + by^2 + cz^2 + bct^2.$$

Theorem 3.47 implies that the quadratic form $Q' = -bx^2 - cy^2 - z^2$ has a null vector over all K_v since Q has a null vector over all K_v. Hence Q' has a null vector over K from the $n = 3$ case of the Hasse–Minkowski theorem over K. Using Theorem 3.47 again gives a null vector for Q over K.

Case 2: Let $\operatorname{disc}(Q) = \Delta$ be a non-square in $K^\times$. Let $E = K(\sqrt{\Delta})$ which is also a number field. We show that Q has a null vector over E and then use this to find a null vector over K. As in Case 1, Theorem 3.47 says that $-bx^2 - cy^2 - z^2$ has a null vector over every E_w since Q has a null vector over all K_v and any E_w contains a K_v. Then by the $n = 3$ case of Hasse–Minkowski *over* E there is a null vector for Q' (and thus Q) over E.

Write $E = K + K\sqrt{\Delta}$. Let $V = K^4$ and $W = E^4 = V + V\sqrt{\Delta}$. Let $v_1 + v_2\sqrt{\Delta}$ with $v_i \in V$ be a null vector for Q over E. Then at least one of v_1 and v_2 is non-zero. Furthermore,

$$\begin{aligned} 0 &= Q(v_1 + v_2\sqrt{\Delta}) \\ &= Q(v_1) + Q(v_2\sqrt{\Delta}) + 2\mathrm{B}(v_1, v_2\sqrt{\Delta}) \\ &= Q(v_1) + \Delta Q(v_2) + 2\mathrm{B}(v_1, v_2)\sqrt{\Delta} \end{aligned}$$

where $\mathrm{B}(v_1, v_2)$ is the bilinear form associated to Q. Notice that $Q(v_1)$, $Q(v_2)$, and $\mathrm{B}(v_1, v_2)$ are in K. So $Q(v_1) = -\Delta Q(v_2)$ and $\mathrm{B}(v_1, v_2) = 0$. Thus $Q(v_1) = 0$ if and only if $Q(v_2) = 0$.

Suppose $Q(v_1) = 0$ or $Q(v_2) = 0$. Then either v_1 or v_2 is a null vector for Q over K since at least one of them is non-zero.

Now suppose $Q(v_1) \neq 0$ and $Q(v_2) \neq 0$. Write $a_i = Q(v_i)$ for $i = 1, 2$. So we have $a_1 = -\Delta a_2$. Build an orthogonal basis of V using v_1 and v_2 as the first two vectors in the basis (recall that $\mathrm{B}(v_1, v_2) = 0$). In this basis, $V = Kv_1 \perp Kv_2 \perp U$ and Q is $a_1y_1^2 + a_2y_2^2 + Q''$ with (U, Q'') a two-dimensional quadratic space which is necessarily non-degenerate. Since the discriminant is well-defined up to squares,

$$\Delta \sim a_1a_2 \operatorname{disc}(Q'') = -\Delta a_2^2 \operatorname{disc}(Q'').$$

Thus $\operatorname{disc}(Q'') \sim -1$. This means U is a hyperbolic plane over K by Theorem 2.60. In particular, Q'' has a null vector over K, so Q does as well. □

Remark 4.8. The proof of Theorem 4.7 comes from [9].

4.3 $n = 4$

The proof of the $n = 4$ case of the Hasse–Minkowski theorem over $\mathbb{Q}$ requires the following two theorems. The first is stated without proof.

Theorem 4.9 (Dirichlet)**.** *For a and m in $\mathbb{Z}$ with $(a, m) = 1$, there are infinitely many primes p such that $p \equiv a \bmod m$.*

Recall Definition 3.44 of the Hilbert symbol. For the remainder of this chapter denote $(a, b)_{\mathbb{Q}_v}$ as $(a, b)_v$ with the convention that $(a, b)_\infty = (a, b)_{\mathbb{R}}$.

Theorem 4.10. *Let $a, b \in \mathbb{Q}^\times$.*

1) *There are finitely many v such that $(a, b)_v = -1$.*

2) $\prod_v (a, b)_v = 1$. *In other words, $\#\{v : (a, b)_v = -1\}$ is even.*

Proof. 1) We already know this from Theorem 3.37. More precisely, this means that the only v for which $(a, b)_v$ might be -1 are ∞, 2 and the odd primes p at which a or b is not in $\mathbb{Z}_p^\times$.

2) From the bimultiplicativity and symmetry of the Hilbert symbol (Corollary 3.50), it is enough to show that (2) is true when a and b are -1 or prime. Since $(a, a)_v = (a, -1)_v$, we do not have to check the case when a and b are the same prime number.

Suppose $a = b = -1$. Then $(-1, -1)_v = 1$ for $v \neq 2, \infty$ while

$$(-1, -1)_\infty = (-1, -1)_2 = -1$$

from Example 3.51.

If $a = -1$ and $b = 2$, then $(-1, 2)_v = 1$ for all v since $2 = 1^2 + 1^2$.

Recall the formula for the Hilbert symbol over a local field K with odd residue field characteristic:

$$(a, b)_K = (\pi^m \varepsilon, \pi^n \delta)_K = (-1)^{mn(q-1)/2} \chi(\bar{\varepsilon})^n \chi(\bar{\delta})^m \tag{4.2}$$

where π is a uniformizer, ε and δ are units, q is the size of the residue field and $\chi : (\mathcal{O}_K/\mathfrak{m}_K)^\times \to \{\pm 1\}$ is the quadratic character on the residue field. We use this formula in the remaining cases for $\mathbb{Q}_p$ where p is odd. The quadratic character on the residue field of $\mathbb{Q}_p$ is the Legendre symbol $\left(\frac{\cdot}{p}\right)$.

Suppose $a = -1$ and $b = l$ for an odd prime l. Then $(-1, l)_v = 1$ for odd $v \neq l$ and $v = \infty$. Over $\mathbb{Q}_l$, $(-1, l)_l = \left(\frac{-1}{l}\right)$ by (4.2). Over $\mathbb{Q}_2$, the equation $-x^2 + ly^2 - z^2 = 0$ has a non-trivial solution (by Corollary 3.42) if and only if there is a primitive solution to the congruence

$$-x^2 + ly^2 - z^2 \equiv 0 \bmod 8.$$

If l is 1 mod 4 (so $l \equiv 1$ or 5 mod 8) then there clearly is a primitive solution. If l is 3 mod 4, then a tedious check shows that there are no primitive solutions. Hence we have $(-1, l)_2 = (-1)^{(l-1)/2}$. So

$$\prod_v (-1, l)_v = (-1, l)_l (-1, l)_2 = \left(\frac{-1}{l}\right)(-1)^{(l-1)/2},$$

which is 1 if and only if the first supplementary law of quadratic reciprocity is true.

Now suppose $a = 2$ and $b = l$ where l is an odd prime. As before, $(2, l)_v = 1$ for all odd primes $v \neq l$ and $v = \infty$. Also, $(2, l)_l = \left(\frac{2}{l}\right)$ using (4.2). Over $\mathbb{Q}_2$, we consider the congruence

$$2x^2 + ly^2 - z^2 \equiv 0 \bmod 8.$$

If l is 1 or 7 mod 8, then it is clear that $(2, l)_2 = 1$ since the congruence becomes

$$2x^2 + y^2 - z^2 \equiv 0 \bmod 8$$

or

$$2x^2 - y^2 - z^2 \equiv 0 \bmod 8.$$

If $l \equiv 3$ or 5 mod 8 a check shows there are no primitive solutions to the congruence. Thus $(2, l)_2 = (-1)^{(l^2-1)/8}$, so

$$\prod_v (2, l)_v = (2, l)_l (2, l)_2 = \left(\frac{2}{l}\right)(-1)^{(l^2-1)/8},$$

which is 1 if and only if the second supplementary law of quadratic reciprocity is true.

Finally, suppose $a = l$ and $b = l'$ for distinct odd primes l, l'. Then $(a, b)_v = 1$ for all v other than l, l', and 2. The Hilbert symbol formula (4.2) shows $(l, l')_l = \left(\frac{l'}{l}\right)$ and $(l', l)_{l'} = \left(\frac{l}{l'}\right)$. In a similar manner as before, the congruence

$$lx^2 + l'y^2 - z^2 \equiv 0 \bmod 8$$

has primitive solutions when at least one of l, l' is 1 mod 4, but has no primitive solutions when both l and l' are 3 mod 4. So

$$(l,l')_2 = \begin{cases} 1, & \text{if } l \text{ or } l' \equiv 1 \bmod 4 \\ -1, & \text{if } l, l' \equiv 3 \bmod 4 \end{cases} = (-1)^{\frac{l-1}{2}\frac{l'-1}{2}}.$$

Thus we have

$$\prod_v (l,l')_v = (l,l')_l (l,l')_{l'} (l,l')_2 = \left(\frac{l'}{l}\right)\left(\frac{l}{l'}\right)(-1)^{\frac{l-1}{2}\frac{l'-1}{2}},$$

which is 1 if and only if the main law of quadratic reciprocity holds. □

Remark 4.11. The second statement in Theorem 4.10 is known as Hilbert reciprocity. As one can see from the proof of this theorem, the non-trivial cases of Hilbert reciprocity are equivalent to quadratic reciprocity, with no special role for 2 or positivity as in the classical statement of quadratic reciprocity. All primes (including ∞) are on an equal footing. Furthermore, Hilbert reciprocity generalizes to all global fields.

Remark 4.12. Hilbert reciprocity says that $(a,b)_v = -1$ for an even number of v. For example, recall that the equation $-x^2 - y^2 - z^2 = 0$ in the proof of the $n = 3$ case of Hasse–Minkowski has a non-trivial solution in $\mathbb{Q}_v$ for all v except $v = 2$ and $v = \infty$.

Corollary 4.13. *For any non-degenerate quadratic form Q over $\mathbb{Q}$, $c_v(Q) = -1$ for only finitely many v and $\prod_v c_v(Q) = 1$.*

Proof. Diagonalize Q and use Hilbert reciprocity. □

Corollary 4.14. *Let $a, b, c \in \mathbb{Q}^\times$. If $ax^2 + by^2 + cz^2 = 0$ has non-trivial solutions in $\mathbb{Q}_v$ for all v except maybe one v_0, then it also has a non-trivial solution in $\mathbb{Q}_{v_0}$.*

Proof. Having a non-trivial solution to $ax^2+by^2+cz^2 = 0$ in $\mathbb{Q}_v$ is equivalent to having a non-trivial solution to $-\frac{a}{c}x^2 - \frac{b}{c}y^2 - z^2 = 0$ in $\mathbb{Q}_v$, which is equivalent to $(-\frac{a}{c}, -\frac{b}{c})_v = 1$. By assumption, $(-\frac{a}{c}, -\frac{b}{c})_v = 1$ for all $v \neq v_0$. So Hilbert reciprocity implies $(-\frac{a}{c}, -\frac{b}{c})_{v_0} = 1$ too. □

When Legendre first proved the $n = 3$ case of the Hasse–Minkowski theorem over $\mathbb{Q}$, he used prime power congruences instead of p-adic numbers. Interestingly, his proof ignores congruences modulo powers of 2 (i.e., the

2-adic case). It is also possible to prove the $n = 3$ case without paying attention to the existence of real solutions; see [2, pp. 79 – 81] for such a proof. Corollary 4.14 explains why this is possible.

We are now ready to prove the $n = 4$ case of the Hasse–Minkowski theorem over $\mathbb{Q}$.

Theorem 4.15. *For non-zero a, b, c, d in $\mathbb{Q}$, the equation*

$$ax^2 + by^2 + cz^2 + dt^2 = 0$$

has a non-trivial solution over $\mathbb{Q}$ if and only if it has a non-trivial solution over all $\mathbb{Q}_v$.

Proof. It suffices to show that non-trivial solvability over all $\mathbb{Q}_v$ implies non-trivial solvability over $\mathbb{Q}$ (the converse is trivial).

By scaling we may assume a, b, c and d are in $\mathbb{Z}$. Since there is a solution over $\mathbb{R}$, without loss of generality assume that $a > 0$ and $d < 0$. Let $Q_1(x, y) = ax^2 + by^2$ and $Q_2(z, t) = -cz^2 - dt^2$ so

$$ax^2 + by^2 + cz^2 + dt^2 = Q_1(x, y) - Q_2(z, t). \tag{4.3}$$

By Corollary 2.45, Q_1 and Q_2 take on a common value $\alpha_p \in \mathbb{Q}_p^\times$ for each prime p. The goal is to use these common values to find our desired non-trivial solution over $\mathbb{Q}$. (The real solvability was used to arrange Q_1 and Q_2 to each have a positive coefficient.)

By scaling, we may assume α_p is in $\mathbb{Z}_p^\times$ or $p\mathbb{Z}_p^\times$. Choose a positive integer r such that $r \equiv \alpha_p \bmod p^2\mathbb{Z}_p$ for odd $p|abcd$ and $r \equiv \alpha_2 \bmod 16\mathbb{Z}_2$. Then $\mathrm{ord}_p(r) = \mathrm{ord}_p(\alpha_p)$ for all $p|2abcd$, so r/α_p is in $\mathbb{Z}_p^\times$ with

$$\frac{r}{\alpha_p} \equiv 1 \bmod p\mathbb{Z}_p \text{ for odd } p|abcd$$

and

$$\frac{r}{\alpha_2} \equiv 1 \bmod 8\mathbb{Z}_2.$$

Thus r/α_p is in $\mathbb{Q}_p^{\times^2}$ for all $p|2abcd$. So Q_1 and Q_2 both take on the value r in these $\mathbb{Q}_p$. Put another way, the 3-dimensional quadratic forms

$$Q_1' = Q_1(x, y) - rw_1^2$$

and

$$Q_2' = Q_2(z, t) - rw_2^2$$

have null vectors in $\mathbb{Q}_p^3$ for all $p|2abcd$. Furthermore, Q_1' and Q_2' have rational coefficients so if we can find non-trivial solutions to $Q_1' = 0$ and $Q_2' = 0$ over every $\mathbb{Q}_v$, then we can apply Hasse–Minkowski for $n = 3$ to get non-trivial solutions to $Q_1' = 0$ and $Q_2' = 0$ over $\mathbb{Q}$.

The equations $Q_1' = 0$ and $Q_2' = 0$ have non-trivial solutions over $\mathbb{R}$ since $a, r, -d > 0$. For p not dividing $2abcdr$ there are always non-trivial solutions over $\mathbb{Q}_p$ to these equations by Theorem 3.37. Are there non-trivial solutions over $\mathbb{Q}_p$ for $p|r$ but not dividing $2abcd$? Rather than answering this question, we show that a positive integer r' can be found such that r' fits all of the conditions imposed on r, and every prime factor of r' except one divides $2abcd$.

Notice that r only matters through its positivity and its congruence class modulo m^2, where

$$m = 4 \prod_{\substack{p|abcd \\ p \neq 2}} p,$$

and possibly $(r, m^2) \neq 1$ since some of the α_p's may have $\mathrm{ord}_p(\alpha_p) > 0$. Let $\delta = (r, m^2)$. Then $(\frac{r}{\delta}, \frac{m^2}{\delta}) = 1$. Dirichlet's theorem says that there are infinitely many primes l such that

$$l \equiv \frac{r}{\delta} \bmod \frac{m^2}{\delta}.$$

Pick such an l that does not divide $2abcd$ and set $r' = l\delta$. Then $r' > 0$ and $r' \equiv r \bmod m^2$, so r' fits all of the conditions imposed on r and every prime dividing r' also divides $2abcd$ except l. Thus, replacing r with r' in Q_1' and Q_2', there are non-trivial solutions to $Q_1' = 0$ and $Q_2' = 0$ over every $\mathbb{Q}_v$ with the possible exception of $\mathbb{Q}_l$. Corollary 4.14 implies that there are also non-trivial solutions to these equations over $\mathbb{Q}_l$.

Using the $n = 3$ case of Hasse–Minkowski, there is a non-trivial solution to $Q_1' = 0$ and $Q_2' = 0$ over $\mathbb{Q}$. Thus there are solutions to $Q_1(x, y) = r'$ and $Q_2(z, t) = r'$ over $\mathbb{Q}$. So by (4.3) there is a non-trivial solution to $ax^2 + by^2 + cz^2 + dt^2 = 0$ over $\mathbb{Q}$. □

4.4 $n \geq 5$

The proof for $n \geq 5$ will follow by induction.

Theorem 4.16. *Let $n \geq 5$. Let $Q(x_1, \ldots, x_n) = a_1x_1^2 + \cdots + a_nx_n^2$ with $a_i \in \mathbb{Q}^\times$. Then $Q(x_1, \ldots, x_n) = 0$ has a non-trivial solution in $\mathbb{Q}$ if and only if it has a non-trivial solution over all $\mathbb{Q}_v$.*

Proof. We will prove the "if" direction; the other direction is trivial. Write $Q = Q_1 - Q_2$ where

$$Q_1(x_1, x_2) = a_1x_1^2 + a_2x_2^2$$

and

$$Q_2(x_3, \ldots, x_n) = -a_3x_3^2 - \cdots - a_nx_n^2.$$

Let S be the set consisting of of $v = 2$, $v = \infty$ and v such that not every $a_i \in \mathbb{Z}_v^\times$ for $i \geq 3$. For all $v \in S$, Q_1 and Q_2 represent some common non-zero α_v over $\mathbb{Q}_v$ since Q has a null vector over $\mathbb{Q}_v$ (Corollary 2.45). That is,

$$Q_1(x_{1,v}, x_{2,v}) = \alpha_v = Q_2(x_{3,v}, \ldots, x_{n,v})$$

for $x_{i,v} \in \mathbb{Q}_v$. The set of non-zero squares $\mathbb{Q}_v^{\times^2}$ is open so the coset of $\mathbb{Q}_v^{\times^2}$ containing α_v is an open set. The quadratic form Q_1 is continuous (it is a polynomial) so the inverse image of the coset containing α_v is an open set A_v in $\mathbb{Q}_v \times \mathbb{Q}_v$. By the approximation theorem (Theorem 3.9) there are $x_1, x_2 \in \mathbb{Q}$ such that $(x_1, x_2) \in A_v$ for all $v \in S$. Thus $a := Q_1(x_1, x_2)$ is in $\mathbb{Q}$ and $a/\alpha_v \in \mathbb{Q}_v^{\times^2}$ for all $v \in S$. Consider the quadratic form $Q' = at^2 - Q_2$. There is a non-trivial solution to $Q' = 0$ over every $\mathbb{Q}_v$ for $v \in S$ since $a/\alpha_v \in \mathbb{Q}_v^{\times^2}$ for all $v \in S$. Furthermore, the equation $Q' = 0$ has a non-trivial solution over every $\mathbb{Q}_v$ where v is not in S since Q_2 is universal over $\mathbb{Q}_v$ by Theorem 3.37 ($n-2$ is at least 3). So there is a non-trivial solution to $Q' = 0$ over $\mathbb{Q}$ by the induction hypothesis since Q' is an $(n-1)$-dimensional quadratic form. This means the equation $Q_2 = a$ has a solution over $\mathbb{Q}$. We now have solutions over $\mathbb{Q}$ to $Q_1 = a$ and $Q_2 = a$ so

$$Q = Q_1 - Q_2 = 0$$

has a non-trivial solution over $\mathbb{Q}$. □

4.5 Applications

An interesting consequence of Theorem 3.63 is that every indefinite (not all terms in a diagonalization have the same sign) quadratic form of dimension greater than or equal to 5 over $\mathbb{Q}$ has a null vector. As a result, when

Q is a four-dimensional quadratic form over $\mathbb{Q}$, the equation $Q = r$ for $r \in \mathbb{Q}^\times$ is solvable if and only if the five-dimensional quadratic form $Q - rx^2$ is indefinite. For example, the quadratic form $x^2+y^2+z^2-7t^2$ from Example 2.44 is universal over $\mathbb{Q}$ since for any $r \in \mathbb{Q}^\times$,

$$x^2 + y^2 + z^2 - 7t^2 - ru^2$$

is indefinite.

We return to the theorems in Chapter 1 about integers as sums of integer squares. We restate Theorem 1.3 and prove it.

Theorem 4.17 (Legendre)**.** *Write a positive integer n in the form $4^a n'$ with $a \geq 0$ and $n' \not\equiv 0 \bmod 4$. Then n is a sum of three integer squares if and only if $n' \not\equiv 7 \bmod 8$.*

Proof. If $n = x^2 + y^2 + z^2$ for some $x, y, z \in \mathbb{Z}$ then $n' = x'^2 + y'^2 + z'^2$ for some $x', y', z' \in \mathbb{Q}$ (just scale n by $\frac{1}{4^a}$). By Theorem 1.6, n' is a sum of three integer squares. So $n' \not\equiv 7 \bmod 8$ (the squares modulo 8 are 0, 1, and 4).

Conversely, suppose $n' \not\equiv 7 \bmod 8$. Then n' is 1, 2, 3, 5, or 6 modulo 8 because 4 does not divide n'. We show that n' is a sum of three rational squares using the Hasse–Minkowski theorem.

For all odd primes p, $x^2 + y^2 + z^2$ has a null vector over $\mathbb{Q}_p$ by Theorem 3.37 and thus it is universal over $\mathbb{Q}_p$ by Theorem 2.39. Over $\mathbb{R}$, $x^2 + y^2 + z^2$ represents n' since $n' > 0$. Over $\mathbb{Q}_2$, we show that n' is $x^2 + y^2 + z^2 \bmod 8$:

- if $n' \equiv 1 \bmod 8$, use $(x, y, z) = (1, 0, 0)$,
- if $n' \equiv 2 \bmod 8$, use $(x, y, z) = (1, 1, 0)$,
- if $n' \equiv 3 \bmod 8$, use $(x, y, z) = (1, 1, 1)$,
- if $n' \equiv 5 \bmod 8$, use $(x, y, z) = (1, 2, 0)$,
- if $n' \equiv 6 \bmod 8$, use $(x, y, z) = (1, 1, 2)$.

Therefore $x^2 + y^2 + z^2 - n't^2$ has a primitive solution modulo 8 and thus a 2-adic solution by Corollary 3.42. So $x^2 + y^2 + z^2$ represents n' over $\mathbb{Q}_2$. Hence n' is a sum of three rational squares. Then Theorem 1.6 implies that n' is a sum of three integer squares and therefore n is a sum of three integer squares too. □

We now restate and prove Theorem 1.4 as a consequence of Theorem 4.17.

Theorem 4.18 (Lagrange)**.** *Every positive integer is a sum of four integer squares.*

Proof. Let $n \in \mathbb{Z}^+$. Write $n = 4^a n'$ with 4 not dividing n'. If $n' \not\equiv 7 \bmod 8$ then it (and consequently n) is a sum of three integer squares by Theorem 4.17. If n' is 7 modulo 8 then $n' - 1$ is a sum of three integer squares. So $n' = x_0^2 + y_0^2 + z_0^2 + 1$ for some $x_0, y_0, z_0 \in \mathbb{Z}$ and thus n is a sum of four integer squares. □

The following theorem is closely related to the Hasse–Minkowski theorem.

Theorem 4.19. *Two non-degenerate quadratic forms with rational coefficients are equivalent over $\mathbb{Q}$ if and only if they are equivalent over every $\mathbb{Q}_v$.*

Proof. As in the proof of the Hasse–Minkowski theorem it is clear that only one direction needs proving. Suppose that two non-degenerate quadratic forms with rational coefficients Q_1 and Q_2 are equivalent over all $\mathbb{Q}_v$. We will show that $Q_1 \cong Q_2$ over $\mathbb{Q}$ by inducting on the dimension n.

Let $n = 1$. Then $Q_1 = ax^2$ and $Q_2 = bx^2$ for $a, b \in \mathbb{Q}^\times$. Since $Q_1 \cong Q_2$ over $\mathbb{Q}_v$, $\frac{a}{b} \in \mathbb{Q}_v^{\times^2}$. So Theorem 4.3 implies $\frac{a}{b} \in \mathbb{Q}^{\times^2}$. Hence $Q_1 \cong Q_2$ over $\mathbb{Q}$.

Now let $n > 1$ and assume that the theorem holds for all lower dimensional non-degenerate quadratic forms with rational coefficients. Let w be a vector over $\mathbb{Q}$ such that $Q_1(w) = a$ where $a \in \mathbb{Q}^\times$. Then $Q_2(w_v) = a$ for some vector w_v over each $\mathbb{Q}_v$ since $Q_1 \cong Q_2$ over all $\mathbb{Q}_v$. So $Q_2(w') = a$ for some vector w' over $\mathbb{Q}$ by the Hasse–Minkowski theorem. Thus,

$$Q_1 \cong ax_1^2 + Q_1'$$

and

$$Q_2 \cong ay_1^2 + Q_2'$$

over $\mathbb{Q}$ where Q_1' and Q_2' are $(n-1)$-dimensional quadratic forms. Since $Q_1 \cong Q_2$ over every $\mathbb{Q}_v$ and $ax_1^2 \cong ay_1^2$ over $\mathbb{Q}$ (and every $\mathbb{Q}_v$), Witt cancellation says that $Q_1' \cong Q_2'$ over every $\mathbb{Q}_v$. So $Q_1' \cong Q_2'$ over $\mathbb{Q}$ by the induction hypothesis. Hence $Q_1 \cong Q_2$ over $\mathbb{Q}$. □

Remark 4.20. Theorem 4.19 is referred to as the weak Hasse–Minkowski theorem. Hasse originally proved it without using Witt cancellation or the Hasse–Minkowski theorem. The usual Hasse–Minkowski theorem is called (in comparison) the strong Hasse–Minkowski theorem. Furthermore, the strong

Hasse–Minkowski theorem can be derived from the weak Hasse–Minkowski theorem.

Example 4.21. Recall the trace forms over $\mathbb{Q}$

$$4x^2 - 36z^2 + 192t^2 + 48xt + 48yz - 72yt \tag{4.4}$$

and

$$4x^2 - 12y^2 + 48z^2 - 168t^2 - 24xz + 24xt + 24yz + 96yt - 240zt \tag{4.5}$$

from Example 1.13 and Example 1.14. We now show them to be inequivalent over $\mathbb{Q}_{11}$ which means they are inequivalent over $\mathbb{Q}$. As a result, the fields K and L from Example 1.13 and Example 1.14 are not isomorphic.

The trace form (4.4) is equivalent to the diagonal quadratic form

$$Q_1(x, y, z, t) = x^2 - y^2 + z^2 - 33t^2$$

by Example 2.55 whereas (4.5) is equivalent (by work omitted here) to

$$Q_2(x, y, z, t) = x^2 - 10y^2 + 165z^2 - 2t^2.$$

The dimensions of Q_1 and Q_2 are both 4 and

$$\operatorname{disc}(Q_1) = (-1)(-33) = 33$$

while

$$\operatorname{disc}(Q_2) = -2 \cdot 5 \cdot 3 \cdot 5 \cdot \cdot 11 \cdot (-2) \sim 33,$$

so we use Hasse invariants to show $Q_1 \ncong Q_2$ over $\mathbb{Q}$. For any v,

$$c_v(Q_1) = (-1, -33)_v = (-1, -3)_v(-1, 11)_v$$

and (after a lot of simplifying)

$$c_v(Q_2) = (-1, -1)_v(5, 2 \cdot 3 \cdot 11)_v.$$

For $v = 11$,

$$c_{11}(Q_1) = (-1, 11)_{11} = -1$$

while

$$c_{11}(Q_2) = \left(\frac{5}{11}\right) = \left(\frac{11}{5}\right) = 1.$$

Thus Q_1 and Q_2 are inequivalent over $\mathbb{Q}$. The reader can check that $c_v(Q_1)$ and $c_v(Q_2)$ are equal to 1 when $v \neq 2$, 3, 11, or ∞ and that they are equal to -1 when $v = \infty$ and 3. Therefore, by Corollary 4.14, $c_2(Q_1) = -1$ and $c_2(Q_2) = 1$ (or use the 2-adic formula in [8, p. 20]).

We conclude this chapter with some necessary and sufficient conditions for deciding when 2-dimensional and 3-dimensional quadratic forms over $\mathbb{Q}$ take on particular rational values.

Theorem 4.22. *Let Q be a 2-dimensional non-degenerate quadratic form over $\mathbb{Q}$. Then Q takes on the value b in $\mathbb{Q}^\times$ if and only if*

$$(b, -\operatorname{disc}(Q))_v = c_v(Q) \text{ for all } v. \tag{4.6}$$

Proof. Use Lemma 3.58 and the Hasse–Minkowski theorem. □

Remark 4.23. Both sides of (4.6) are -1 at only finitely many places (Theorem 4.10 and Corollary 4.13). So verifying (4.6) is only a finite calculation.

Theorem 4.24. *Let Q be a 3-dimensional non-degenerate quadratic form over $\mathbb{Q}$. Then Q takes on a value b in $\mathbb{Q}^\times$ if and only if for each v at least one of the following conditions holds*:

1) $-b \operatorname{disc}(Q) \notin \mathbb{Q}_v^{\times^2}$,

2) $c_v(Q) = (-1,-1)_v(\operatorname{disc}(Q), -b)_v$.

Proof. Let $Q' = Q - bx^2$. We already know that $Q = b$ over $\mathbb{Q}$ if and only if Q' has a null vector over $\mathbb{Q}$. This last statement is equivalent to Q' having a null vector over all $\mathbb{Q}_v$ by the Hasse–Minkowski theorem. But Q' has a null vector over all $\mathbb{Q}_v$ if and only if Theorem 3.61 is satisfied over all $\mathbb{Q}_v$. Since

$$c_v(Q') = c_v(Q)(\operatorname{disc}(Q), -b)_v,$$

the condition $c_v(Q') = (-1,-1)_v$ in Theorem 3.61 is the same as

$$c_v(Q) = (-1,-1)_v(\operatorname{disc}(Q), -b)_v.$$

□

Chapter 5

Hasse–Minkowski Over $\mathbb{F}(T)$

The motivation for this chapter is to show the similarities between $\mathbb{Q}$ and $\mathbb{F}(T)$ in the context of the Hasse–Minkowski theorem. The reader should be careful, however, because in some instances there are differences that cannot be overlooked for our proofs.

The statement of Hasse–Minkowski over $\mathbb{F}(T)$ that we will prove is the following:

Theorem 5.1. *Let $Q(x_1, \ldots, x_n)$ be a non-degenerate quadratic form over $\mathbb{F}(T)$. The equation $Q(x_1, \ldots, x_n) = 0$ is non-trivially solvable over $\mathbb{F}(T)$ if and only if it is non-trivially solvable over every $\mathbb{F}(T)_v$.*

As in the rational case, a more general version of the Hasse–Minkowski theorem over $\mathbb{F}(T)$ in the spirit of Theorem 1.16 follows from Theorem 5.1 in a purely algebraic manner.

Since $\mathbb{F}$ has odd characteristic and $\mathbb{F} \subset \mathbb{F}(T)_v$ for all v, the residue field of $\mathbb{F}(T)_v$ has the same odd characteristic as $\mathbb{F}$. This means we will not run into the complications that arose from $\mathbb{Q}_2$. Other issues will arise, however, from $\mathbb{F}(T)_\infty$.

Just like the rational case, we may assume Q in Theorem 5.1 is in diagonal form.

5.1 $n = 2$

Just as in the rational case, we must show that for any $f, g \in \mathbb{F}(T)^\times$ if the equation $fx^2 + gy^2 = 0$ has a non-trivial solution over all $\mathbb{F}(T)_v$ then that there is a non-trivial solution to $fx^2 + gy^2 = 0$ over $\mathbb{F}(T)$.

Use Lemma 4.2 to reduce the proof to the following theorem which is the analogue of the square theorem over $\mathbb{Q}$.

Theorem 5.2. *Let $h \in \mathbb{F}(T)^\times$. Then h is a square in $\mathbb{F}(T)$ if and only if h is a square in all $\mathbb{F}(T)_v$.*

Proof. Only one direction needs proving since $\mathbb{F}(T) \subset \mathbb{F}(T)_v$. Factor h as

$$h = c\pi_1^{e_1}\pi_2^{e_2}\dots\pi_r^{e_r}$$

where the π_i's are distinct monic irreducibles and $c \in \mathbb{F}^\times$. Suppose h is a square in all $\mathbb{F}(T)_v$. Then $\operatorname{ord}_{\pi_i}(h) = e_i$ is even for every π_i so c is a square in every $\mathbb{F}(T)_v$. Taking $v = \infty$, c is a square in $\mathbb{F}$ since c is a square in $\mathbb{F}(T)_\infty$ and the residue field of $\mathbb{F}(T)_\infty$ is $\mathbb{F}$. Hence h is a square in $\mathbb{F}(T)$. □

5.2 $n = 3$

In this case we will show that if there is a non-trivial solution over every $\mathbb{F}(T)_v$ to $fx^2 + gy^2 + hz^2 = 0$ where f, g, and h are in $\mathbb{F}(T)^\times$, then there is a non-trivial solution over $\mathbb{F}(T)$ to $fx^2 + gy^2 + hz^2 = 0$.

We make the same reductions as over $\mathbb{Q}$. We may assume the coefficients f, g, and h to be in the ring $\mathbb{F}[T]$ since scaling by a common denominator of the coefficients does not affect the existence of a non-trivial solution. If any of the coefficients have square factors, then by a linear change of variables we can get f, g, and h to be square-free. Since $fx^2 + gy^2 + hz^2 = 0$ has a non-trivial solution if and only if $fx^2 + gy^2 = -hz^2$ has a non-trivial solution, we can work with the equation $fx^2 + gy^2 = z^2$ where f and g are square-free and are in $\mathbb{F}[T]$ (scale by $-h$, then use a linear change of variables). Instead of inducting on the sum of the absolute values of the coefficients, we will induct on the sum of the degrees of the coefficients.

Theorem 5.3. *Let f and g in $\mathbb{F}[T]$ be square-free. If $fx^2 + gy^2 = z^2$ has a non-trivial solution over every $\mathbb{F}(T)_v$, then it has a non-trivial solution over $\mathbb{F}(T)$.*

Proof. We will induct on $\deg f + \deg g$. Let $\deg f + \deg g = 0$. Then we have $\deg f = \deg g = 0$. So the equations that fit these conditions are $ax^2 + by^2 = z^2$ where $a, b \in \mathbb{F}^\times$. By Lemma 2.74 there is a solution (x_0, y_0) to $ax^2 + by^2 = 1$ over $\mathbb{F}$. Thus we have found a non-trivial solution $(x_0, y_0, 1)$ over $\mathbb{F}$!

Now assume that $\deg f + \deg g > 0$ and that the theorem is true for all smaller values. Without loss of generality assume that $\deg f \leq \deg g$, so $\deg g > 0$.

The argument for why f is a square modulo g is identical to the same assertion over $\mathbb{Q}$. So write $f \equiv h^2 \bmod g$. Without loss of generality, we may assume $h = 0$ or $\deg h < \deg g$. Now write $f + gk = h^2$ for some $k \in \mathbb{F}[T]$. By the hypothesis of the theorem f is square-free (and the case when $f \in \mathbb{F}^{\times 2}$ is trivial) so $k \neq 0$. Let $k = lm^2$ where $l \in \mathbb{F}[T]$ is square-free. Thus dividing $f + glm^2 = h^2$ by $(lm)^2$ gives

$$\begin{aligned} \frac{g}{l} &= \frac{h^2 - f}{(lm)^2} \\ &= \frac{h^2}{(lm)^2} - f\frac{1}{(lm)^2}. \end{aligned} \tag{5.1}$$

By the hypothesis of the theorem $g = x_v^2 - fy_v^2$ in each $\mathbb{F}(T)_v$, so (5.1) shows that l has the same form because $\{x^2 - fy^2 \neq 0 : x, y \in \mathbb{F}(T)_v\}$ is a subgroup of $\mathbb{F}(T)_v^\times$. Hence there are non-trivial solutions to $fx^2 + ly^2 = z^2$ in all $\mathbb{F}(T)_v$.

We now show $\deg f + \deg l < \deg f + \deg g$. That is, $\deg l < \deg g$. If $h = 0$ then $f = -glm^2$, so

$$\deg(glm^2) = \deg f \leq \deg g.$$

Thus $\deg l = 0 < \deg g$. Now suppose $h \neq 0$. Then

$$\begin{aligned} \deg(glm^2) &= \deg g + \deg(lm^2) \\ &= \deg(h^2 - f) \\ &\leq \max(\deg(h^2), \deg f) \\ &\leq \max(2(\deg g - 1), \deg g). \end{aligned}$$

If $\deg g \geq 2$, then the maximum is $2(\deg g - 1)$, so

$$\begin{aligned} \deg l &\leq \deg(lm^2) \\ &\leq \deg(g) - 2 \\ &< \deg g. \end{aligned}$$

If $\deg g = 1$ then $\deg l \leq \deg(lm^2) \leq 0$. So $\deg l = 0$ and consequently $\deg l < \deg g$. Then by the induction hypothesis there is a non-trivial solution to $fx^2 + ly^2 = z^2$ over $\mathbb{F}(T)$. Thus $l = x_0^2 - fy_0^2$ for $x_0, y_0 \in \mathbb{F}(T)$. Again by (5.1), g has the form $x^2 - fy^2$ for some $x, y \in \mathbb{F}(T)$. So $fx^2 + gy^2 = z^2$ has a non-trivial solution over $\mathbb{F}(T)$. □

5.3 $n = 4$

Similar to Hasse–Minkowski over $\mathbb{Q}$, we use analogues of Dirichlet's theorem and Hilbert reciprocity for $\mathbb{F}(T)$ to prove the $n = 4$ case. The analogue of Dirichlet's theorem is stated without proof.

Theorem 5.4. *For f and g in $\mathbb{F}[T]$ with $(f, g) = 1$ and any integers a and b with $b \neq 0$, there are infinitely many monic irreducibles π such that $\pi \equiv f \bmod g$ and $\deg \pi \equiv a \bmod b$.*

Remark 5.5. Note that in Theorem 5.4 we do not require $(a, b) = 1$. The case we will need is $b = 2$.

Before treating Hilbert reciprocity for $\mathbb{F}(T)$, we state quadratic reciprocity for $\mathbb{F}[T]$. For an irreducible π in $\mathbb{F}[T]$ and $f \not\equiv 0 \bmod \pi$, set $\left(\frac{f}{\pi}\right) = 1$ if f is a square modulo π and $\left(\frac{f}{\pi}\right) = -1$ if f is not a square modulo π.

Theorem 5.6. *Let $q = \#\mathbb{F}$. For distinct monic irreducibles π_1 and π_2 in $\mathbb{F}[T]$,*

$$\left(\frac{\pi_2}{\pi_1}\right) = (-1)^{\frac{d_1 d_2 (q-1)}{2}} \left(\frac{\pi_1}{\pi_2}\right) \tag{5.2}$$

where $d_1 = \deg \pi_1$, $d_2 = \deg \pi_2$ and q is the size of $\mathbb{F}$. For $c \in \mathbb{F}^\times$ and a monic irreducible $\pi \in \mathbb{F}[T]$,

$$\left(\frac{c}{\pi}\right) = c^{\frac{d(q-1)}{2}} \tag{5.3}$$

where $d = \deg \pi$.

Remark 5.7. Equation (5.2) is known as the main law of quadratic reciprocity while (5.3) is called the supplementary law of quadratic reciprocity. Proofs of Theorems 5.4 and 5.6 can be found in [7].

Regarding Hilbert symbols, we make the same convention for $\mathbb{F}(T)$ as we did for $\mathbb{Q}$. In other words, we denote $(f, g)_{\mathbb{F}(T)_v}$ as $(f, g)_v$ for the remainder of the chapter.

Theorem 5.8. *Let $f, g \in \mathbb{F}(T)^\times$.*

1) *There are finitely many v such that $(f, g)_v = -1$.*

2) $\prod_v (f, g)_v = 1$. *In other words, $\#\{v : (f, g)_v = -1\}$ is even.*

Proof. 1) This follows from Theorem 3.37.

2) Analogously to the rational case, it is enough to show that (2) is true when f and g are in $\mathbb{F}^\times$ or distinct monic irreducibles in $\mathbb{F}[T]$. We will again use the formula

$$(\pi^m\varepsilon, \pi^n\delta)_K = (-1)^{mn(q-1)/2}\chi(\bar{\varepsilon})^n\chi(\bar{\delta})^m \tag{5.4}$$

to calculate the Hilbert symbols. (Be aware that q is not always $\#\mathbb{F}$; it is $\#\mathbb{F}^{\deg\pi}$ for $v = \pi$ and is $\#\mathbb{F}$ for $v = \infty$.) When $K = \mathbb{F}(T)_\pi$, with residue field isomorphic to $\mathbb{F}[T]/(\pi)$, the quadratic character χ is the Legendre symbol $\left(\frac{\cdot}{\pi}\right)$. When $K = \mathbb{F}(T)_\infty = \mathbb{F}((\frac{1}{T}))$, with residue field isomorphic to $\mathbb{F}$, χ is the quadratic character on $\mathbb{F}^\times$: $\chi(c) = c^{(q-1)/2}$ for $c \in \mathbb{F}^\times$.

Suppose $f = c_1$ and $g = c_2$ for $c_i \in \mathbb{F}^\times$. Then $(c_1, c_2)_v = 1$ for all v by Theorem 3.37.

Suppose $f = c$ and $g = \pi$ for $c \in \mathbb{F}^\times$ and a monic irreducible π. Then $(c, \pi)_v = 1$ for all v except π and ∞. Furthermore $(c, \pi)_\pi = \left(\frac{c}{\pi}\right)$ by (5.4). If $d = \deg\pi$ is even then π is a square in $\mathbb{F}(T)_\infty$ by Example 3.22, so

$$(c, \pi)_\infty = 1 = c^{d(q-1)/2}$$

since $c^{(q-1)/2} = \pm 1$. If d is odd then $\pi \equiv \frac{1}{T} \bmod \mathbb{F}(T)_\infty^{\times 2}$ so

$$(c, \pi)_\infty = (c, 1/T)_\infty = c^{(q-1)/2}$$

by (5.4). Since d is odd and $c^{(q-1)/2} = \pm 1$, we have $(c, \pi)_\infty = c^{d(q-1)/2}$. This means

$$\prod_v (c, \pi)_v = (c, \pi)_\infty (c, \pi)_\pi = c^{d(q-1)/2}\left(\frac{c}{\pi}\right),$$

which is 1 if and only if the supplementary law of quadratic reciprocity holds.

Lastly, suppose $f = \pi_1$ and $g = \pi_2$ where π_1 and π_2 are distinct monic irreducibles. Again, $(\pi_1, \pi_2)_v = 1$ for all v except π_1, π_2, and ∞. Using the Hilbert symbol formula gives $(\pi_1, \pi_2)_{\pi_1} = (\frac{\pi_2}{\pi_1})$ and $(\pi_1, \pi_2)_{\pi_2} = (\frac{\pi_1}{\pi_2})$. To find $(\pi_1, \pi_2)_\infty$ consider the quadratic form

$$\pi_1 x^2 + \pi_2 y^2 - z^2. \tag{5.5}$$

Suppose that at least one of π_1 and π_2 has even degree. Without loss of generality let $\deg\pi_1$ be even. Then (5.5) is equivalent to $x^2 + \pi_2 y^2 - z^2$ which has (1,0,1) as a null vector so $(\pi_1, \pi_2)_\infty = 1$. If both π_1 and π_2 have odd

degree, then (5.5) is equivalent to $\frac{1}{T}(x^2+y^2)-z^2$. So it has a null vector if and only if $x^2+y^2=0$ has a non-trivial solution in $\mathbb{F}(T)_\infty$ by Theorem 3.34. But $x^2+y^2=0$ has a non-trivial solution in $\mathbb{F}(T)_\infty$ if and only if -1 is a square in $\mathbb{F}^\times$. Thus

$$\begin{aligned}(\pi_1,\pi_2)_\infty &= \begin{cases} 1, & \text{if } \deg\pi_1 \text{ or } \deg\pi_2 \equiv 0 \bmod 2,\\ (-1)^{(q-1)/2}, & \text{if } \deg\pi_1, \deg\pi_2 \equiv 1 \bmod 2\end{cases}\\ &= (-1)^{d_1 d_2 (q-1)/2}\end{aligned}$$

where $d_1=\deg\pi_1$ and $d_2=\deg\pi_2$. So

$$\prod_v (\pi_1,\pi_2)_v = (\pi_1,\pi_2)_\infty(\pi_1,\pi_2)_{\pi_1}(\pi_1,\pi_2)_{\pi_2} = (-1)^{d_1 d_2 (q-1)/2}\left(\frac{\pi_2}{\pi_1}\right)\left(\frac{\pi_1}{\pi_2}\right),$$

which is 1 if and only if the main law of quadratic reciprocity is true. □

Remark 5.9. As in the rational case, the proof of Theorem 5.8 shows Hilbert reciprocity over $\mathbb{F}(T)$ to be equivalent to quadratic reciprocity on $\mathbb{F}[T]$; also, $\prod_v c_v(Q)=1$ for any non-degenerate quadratic form Q over $\mathbb{F}(T)$.

Corollary 5.10. *Let $f,g,h\in\mathbb{F}(T)^\times$. If $fx^2+gy^2+hz^2=0$ has a non-trivial solution over $\mathbb{F}(T)_v$ for all v except maybe one v_0, then it also has a non-trivial solution over $\mathbb{F}(T)_{v_0}$.*

Proof. The proof follows unchanged from that of Corollary 4.14. □

We now proceed with the proof of the $n=4$ case of the Hasse–Minkowksi theorem over $\mathbb{F}(T)$.

Theorem 5.11. *For non-zero f,g,h,k in $\mathbb{F}(T)$, the equation*

$$fx^2+gy^2+hz^2+kt^2=0$$

has a non-trivial solution over $\mathbb{F}(T)$ if and only if it has a non-trivial solution over all $\mathbb{F}(T)_v$.

Proof. As in the rational case, we only need to prove the if direction.

We may assume that f,g,h and k are in $\mathbb{F}[T]$ because scaling does not change the existence of non-trivial solutions. Since there is a non-trivial solution, say (x_0,y_0,z_0,t_0), over $\mathbb{F}(T)_\infty$, the non-archimedean nature of the absolute value implies that at least two of $fx_0^2, gy_0^2, hz_0^2, kt_0^2$ have maximal

size in $\mathbb{F}(T)_\infty$. So assume without loss of generality that fx_0^2 and kt_0^2 have maximal size in $\mathbb{F}(T)_\infty$. That is,

$$|gy_0^2|_\infty, |hz_0^2|_\infty \leq |fx_0^2|_\infty = |kt_0^2|_\infty. \tag{5.6}$$

In particular, $x_0 \neq 0$ and $t_0 \neq 0$. Like the proof for $n = 4$ over $\mathbb{Q}$, let

$$Q_1(x, y) = fx^2 + gy^2$$

and

$$Q_2(z, t) = -hz^2 - kt^2.$$

So Q_1 and Q_2 take on a common value $\alpha_v \in \mathbb{F}(T)_v^\times$ for every v.

For $v = \pi$, we may assume α_π is in $\mathbb{F}[T]_\pi^\times$ or $\pi\mathbb{F}[T]_\pi^\times$. Then choose a polynomial r in $\mathbb{F}[T]$ such that $r \equiv \alpha_\pi \bmod \pi^2\mathbb{F}[T]_\pi$ for π dividing $fghk$. So r/α_π is a non-zero square in $\mathbb{F}(T)_\pi^\times$ for all $\pi|fghk$ by the same reasoning used over $\mathbb{Q}$. Hence the equations

$$Q_1(x, y) - rw_1^2 = 0 \tag{5.7}$$

and

$$Q_2(z, t) - rw_2^2 = 0 \tag{5.8}$$

have non-trivial solutions over $\mathbb{F}(T)_\pi$ for all $\pi|fghk$ and notice that all of the coefficients are in $\mathbb{F}(T)$.

It may not happen that (5.7) and (5.8) have non-trivial solutions over $\mathbb{F}(T)_\infty$. We will find an $\widetilde{r} \in \mathbb{F}[T]$ such that $\widetilde{r} \equiv r \bmod M^2$ where M is the product of all π dividing $fghk$ and

$$\left(\frac{f}{\widetilde{r}}, \frac{g}{\widetilde{r}}\right)_\infty = \left(-\frac{h}{\widetilde{r}}, -\frac{k}{\widetilde{r}}\right)_\infty = 1. \tag{5.9}$$

Then the equations $Q_1(x, y) - \widetilde{r}w_1^2 = 0$ and $Q_2(z, t) - \widetilde{r}w_2^2 = 0$ will both have non-trivial solutions over $\mathbb{F}(T)_\pi$ for all $\pi|fghk$ and over $\mathbb{F}(T)_\infty$.

Let $\widetilde{r} = r + \beta T^s M^2$ for $\beta \in \mathbb{F}^\times$ and s to be determined. By choosing s to be large depending on r and M, the degree and leading coefficient of $\widetilde{r}$ do not depend on r. The leading coefficient of $\widetilde{r}$ is β (since M is monic) and $\deg \widetilde{r} \equiv s \bmod 2$. We will see that (5.9) can be satisfied by appropriate choices of the leading coefficient and degree of $\widetilde{r}$.

Recall the Hilbert symbol formula

$$(\pi^m\varepsilon, \pi^n\delta)_\infty = (-1)^{mn(q-1)/2}\chi(\bar{\varepsilon})^n\chi(\bar{\delta})^m,$$

where $\pi = \frac{1}{T}$, $q = \#\mathbb{F}$ and χ is the quadratic character on $\mathbb{F}^\times$.

Note that $\deg f \equiv \deg k \bmod 2$ since $|fx_0^2|_\infty = |kt_0^2|_\infty \neq 0$. Let s be congruent to $\deg f \bmod 2$, so $f/\widetilde{r}$ and $-k/\widetilde{r}$ in each of the Hilbert symbols in (5.9) have even valuation.

We now consider four cases and show that there is an $\widetilde{r}$ satisfying (5.9) in each case. Suppose $\deg f \equiv \deg g \bmod 2$ and $\deg h \equiv \deg k \bmod 2$. Then (5.9) is satisfied with large $s \equiv \deg f \bmod 2$ and $\beta = 1$ because $f/\widetilde{r}$, $g/\widetilde{r}$, $-h/\widetilde{r}$, $-k/\widetilde{r}$ all have even valuation.

Now suppose $\deg f \equiv \deg g \bmod 2$ but $\deg h \not\equiv \deg k \bmod 2$. For large $s \equiv \deg f \bmod 2$ and any $\beta \in \mathbb{F}^\times$, the first Hilbert symbol in (5.9) is 1 but the second is $\chi(\bar{\delta})$ where χ is the quadratic character on $\mathbb{F}^\times$ and δ is the unit part of $-k/\widetilde{r}$. The reduction of the unit part of $-k/\widetilde{r}$ is the ratio of the leading coefficients of $-k$ and $\widetilde{r}$. Choose β, the leading coefficient of $\widetilde{r}$, to have the same quadratic character as the leading coefficient of $-k$. So $(-h/\widetilde{r}, -k/\widetilde{r})_\infty = 1$ because the ratio of the leading coefficients of $-k$ and $\widetilde{r}$ is a square.

If $\deg f \not\equiv \deg g \bmod 2$ and $\deg h \equiv \deg k \bmod 2$, then by an analogous argument to the previous case (5.9) can be satisfied for suitable $\widetilde{r}$.

Finally, suppose $\deg f \not\equiv \deg g \bmod 2$ and $\deg h \not\equiv \deg k \bmod 2$. We will show that the ratio of the leading coefficients of f and $-k$ is a square and then settle this case. By (5.6) we have that $|gy_0^2|_\infty \leq |fx_0^2|_\infty$. The inequality is strict, however, since $\deg f \not\equiv \deg g \bmod 2$. By a similar argument $|hz_0^2|_\infty < |kt_0^2|_\infty$.

Since the inequality in (5.6) is strict and

$$fx_0^2 + gy_0^2 + hz_0^2 + kt_0^2 = 0,$$

the first non-zero coefficients of the Laurent expansions of fx_0^2 and kt_0^2 in $\mathbb{F}(T)_\infty = \mathbb{F}((\frac{1}{T}))$ must cancel. Let λ_0 and μ_0 be the first non-zero coefficients in the Laurent expansions of x_0 and t_0 respectively. The first non-zero coefficient in a Laurent expansion in $\mathbb{F}(T)_\infty$ of a polynomial in $\mathbb{F}[T]$ is the leading coefficient of the polynomial (see Example 3.22). So

$$(\operatorname{lead} f)\lambda_0^2 + (\operatorname{lead} k)\mu_0^2 = 0$$

in $\mathbb{F}$. This means $-(\operatorname{lead} f)/(\operatorname{lead} k)$ is a square in $\mathbb{F}^\times$. Hence the leading coefficients of f and $-k$ have the same quadratic character. Then choosing $\widetilde{r}$ such that its degree s is large and satisfies $s \equiv \deg f \bmod 2$ and letting

its leading coefficient β have the same quadratic character as $\operatorname{lead} f$ (and consequently $\operatorname{lead}(-k)$) is enough to satisfy (5.9).

We now know that a non-zero $\widetilde{r}$ can be found in $\mathbb{F}[T]$ such that

$$Q_1(x,y) - \widetilde{r}w_1^2 = 0 \tag{5.10}$$

and

$$Q_2(z,t) - \widetilde{r}w_2^2 = 0 \tag{5.11}$$

have non-trivial solutions over $\mathbb{F}(T)_\pi$ for $\pi | fghk$ and over $\mathbb{F}(T)_\infty$. For π not dividing $fghk\widetilde{r}$, (5.10) and (5.11) have non-trivial solutions over $\mathbb{F}(T)_\pi$ by Theorem 3.37. We now find an r' such that (5.10) and (5.11) with r' in place of $\widetilde{r}$ have non-trivial solutions over all $\mathbb{F}(T)_v$ except for possibly one v.

Let δ be the monic $\gcd(\widetilde{r}, M^2)$. Then $(\frac{\widetilde{r}}{\delta}, \frac{M^2}{\delta}) = 1$. So Dirichlet's theorem on $\mathbb{F}[T]$ (Theorem 5.4) shows that for each $c \in \mathbb{F}^\times$ there are infinitely many monic irreducibles $\widetilde{\pi}$ in $\mathbb{F}[T]$ such that

$$\widetilde{\pi} \equiv \frac{\widetilde{r}}{c\delta} \bmod \frac{M^2}{\delta}.$$

Pick such a $\widetilde{\pi}$ not dividing $fghk$. Let $r' = c\widetilde{\pi}\delta$. So $r' \equiv \widetilde{r} \bmod M^2$, $\operatorname{lead} r' = c$ and $\deg r' = \deg \widetilde{\pi} + \deg \delta$. Furthermore, the only π dividing r' and not dividing $fghk$ is $\widetilde{\pi}$. This means that there are non-trivial solutions to

$$Q_1(x,y) - r'w_1^2 = 0 \tag{5.12}$$

and

$$Q_2(z,t) - r'w_2^2 = 0 \tag{5.13}$$

over all $\mathbb{F}(T)_v$ except possibly $v = \infty$ and $v = \widetilde{\pi}$. However, we can guarantee non-trivial solutions over $\mathbb{F}(T)_\infty$ to (5.12) and (5.13) by arranging $r'/\widetilde{r}$ to be a square in $\mathbb{F}(T)_\infty$. This is equivalent to $\deg r' \equiv \deg \widetilde{r} \bmod 2$ and $\chi(\operatorname{lead} r')$ equals $\chi(\operatorname{lead} \widetilde{r})$. In other words, choose c such that $\chi(c) = \chi(\operatorname{lead} \widetilde{r})$ and $\widetilde{\pi}$ such that $\deg \widetilde{\pi} \equiv \deg \widetilde{r} - \deg \delta \bmod 2$. (This is possible by Theorem 5.4.)

The remainder of the proof follows exactly the same argument as the $n = 4$ case over $\mathbb{Q}$, with Corollary 5.10 in place of Corollary 4.14. □

5.4 $n \geq 5$

Theorem 5.12. *Let $n \geq 5$. Let $Q(x_1, \ldots, x_n) = a_1x_1^2 + \cdots + a_nx_n^2$ with $a_i \in \mathbb{F}(T)^\times$. Then $Q(x_1, \ldots, x_n) = 0$ has a non-trivial solution over $\mathbb{F}(T)$ if and only if it has a non-trivial solution over all $\mathbb{F}(T)_v$.*

Proof. Use the same argument as in the $n = 5$ case over $\mathbb{Q}$; no changes are needed. □

Remark 5.13. By Hasse–Minkowski, every non-degenerate 5-dimensional quadratic form over $\mathbb{F}(T)$ has a null vector since every non-degenerate 5-dimensional quadratic form over a local field (of characteristic not 2) has a null vector. Consequently, every non-degenerate 4-dimensional quadratic form over $\mathbb{F}(T)$ is universal.

The following is the weak Hasse–Minkowski theorem over $\mathbb{F}(T)$.

Theorem 5.14. *Two non-degenerate quadratic forms over $\mathbb{F}(T)$ are equivalent over $\mathbb{F}(T)$ if and only if they are equivalent over every $\mathbb{F}(T)_v$.*

Proof. There is no change from the proof of the weak Hasse–Minkowski theorem over $\mathbb{Q}$. □

From the Hasse–Minkowski theorem over $\mathbb{Q}$ and $\mathbb{F}(T)$ as well as Theorem 4.22 and Theorem 4.24 and their analogues over $\mathbb{F}(T)$ we get the following table.

Over $\mathbb{Q}$	Over $\mathbb{F}(T)$
A 2-dimensional non-degenerate quadratic form Q has a value b in $\mathbb{Q}^\times$ if and only if $$(b, -\operatorname{disc}(Q))_v = c_v(Q)$$ for all v.	A 2-dimensional non-degenerate quadratic form Q has a value b in $\mathbb{F}(T)^\times$ if and only if $$(b, -\operatorname{disc}(Q))_v = c_v(Q)$$ for all v.
A 3-dimensional non-degenerate quadratic form Q has a value b in $\mathbb{Q}^\times$ if and only if for all v at least one of the conditions • $-b\operatorname{disc}(Q) \notin \mathbb{Q}_v^{\times^2}$ • $c_v(Q) = (-1,-1)_v(\operatorname{disc}(Q), -b)_v$ holds.	A 3-dimensional non-degenerate quadratic form Q has a value r in $\mathbb{F}(T)^\times$ if and only if for all v at least one of the conditions • $-r\operatorname{disc}(Q) \notin \mathbb{F}(T)_v^{\times^2}$ • $c_v(Q) = (\operatorname{disc}(Q), -r)_v$ holds.
A 4-dimensional non-degenerate quadratic form Q takes on a value b in $\mathbb{Q}^\times$ if and only if the 5-dimensional quadratic form $$Q - bx^2$$ is indefinite.	Every 4-dimensional non-degenerate quadratic form is universal.
A non-degenerate quadratic form of dimension greater than or equal to five has a null vector if and only if it is indefinite.	Any non-degenerate quadratic form of dimension greater than or equal to five has a null vector.

Appendix A

Conics

The aim of this appendix is to tie up a few loose ends from Chapter 1.

Theorem A.1. *If an integer is a sum of two rational squares then it is a sum of two integer squares.*

Proof. We will use geometric ideas involving the circle $x^2 + y^2 = n$. Suppose $P = (p_1, p_2)$ is a rational solution to $x^2 + y^2 = n$ with p_1 or p_2 not in $\mathbb{Z}$. Consider the integer lattice point $M = (m_1, m_2)$ where $|m_i - p_i| \leq 1/2$. The line passing through P and M is not tangent to the circle.

To see why this is true, notice that $0 < |PM|^2 \leq 1/2 < 1$. If the line is tangent to the circle, then the line segments $|PM|$, $|P|$, and $|M|$ form a right triangle with $|M|$ as the hypotenuse. So $|PM|^2 = |M|^2 - |P|^2$ with $|M|^2 - |P|^2 = |M|^2 - n$ in $\mathbb{Z}$. This is a contradiction, however, because $|PM|^2$ is not an integer. Thus the line passing through P and M is not tangent to the circle.

Therefore we have another point P' on the circle and on the line $\overline{PM}$, so

$$\begin{aligned} P' &= P + t(M - P) \\ &= (p_1 + t(m_1 - p_1), p_2 + t(m_2 - p_2)) \end{aligned}$$

for some $t \in \mathbb{R}^\times$. Then defining $v = (m_1 - p_1, m_2 - p_2)$ we get

$$n = P' \cdot P' = (P + tv) \cdot (P + tv)$$

which implies $0 = 2t(v \cdot P) + t^2(v \cdot v)$. So $t = -2(P \cdot v)/(v \cdot v)$. Since P and v have rational coordinates, $t \in \mathbb{Q}^\times$. Thus P' has rational coordinates.

Let d be a positive common denominator of the p_i's. Define

$$\begin{aligned} d' &= d|PM|^2 \\ &= d(v \cdot v) \\ &= d((m_1 - p_1)^2 + (m_2 - p_2)^2) \\ &= d(n - 2(p_1 m_1 + p_2 m_2) + m_1^2 + m_2^2). \end{aligned} \tag{A.1}$$

Notice that $|d'| < |d|$ since $|PM| < 1$, and $d' \in \mathbb{Z}^+$ since $d \in \mathbb{Z}^+$ and $dp_i \in \mathbb{Z}$. So

$$\begin{aligned} t &= -2\frac{P \cdot v}{v \cdot v} \\ &= -2\frac{P \cdot v}{d'/d} \\ &= \frac{-2d(p_1 m_1 + p_2 m_2 - n)}{d'}. \end{aligned} \tag{A.2}$$

Finally, we show that d' is a common denominator for the coordinates of P', namely $p_i + t(m_i - p_i)$. If

$$d'(p_i + t(m_i - p_i)) = d'p_i + d't(m_i - p_i)$$

is an integer, then d' is a common denominator. From equations (A.1) and (A.2) we get

$$\begin{aligned} d't &= d(2n - 2(p_1 m_1 + p_2 m_2)) \\ &= d(2n + d'/d - n - m_1^2 - m_2^2) \\ &= d' + d(n - m_1^2 - m_2^2). \end{aligned}$$

Hence

$$\begin{aligned} d'p_i + d't(m_i - p_i) &= d'p_i + (d' + d(n - m_1^2 - m_2^2))(m_i - p_i) \\ &= d'm_i + d(n - m_1^2 - m_2^2)(m_i - p_i) \end{aligned}$$

which is in $\mathbb{Z}$ because d is a common denominator of the p_i's. So successive repetitions of this algorithm will eventually give an integer solution to $x^2 + y^2 = n$ since the common denominator is reduced each time. □

Theorem A.2. *If an integer is a sum of three rational squares then it is a sum of three integer squares.*

Proof. The proof follows in the same manner as the proof for Theorem A.2, just in more variables. There is one difference, however, in the estimation of the squared distance between the points P and M. For sums of three squares, $0 < |PM|^2 \leq 3/4$, which is still less than one. So we can draw the same conclusion that the line passing through P and M is not tangent to the sphere $x^2 + y^2 + z^2 = n$. □

Remark A.3. The argument used in the proof of Theorem A.2 breaks down, though, for sums of four squares precisely because the estimate of the squared distance between P and M becomes $0 < |PM|^2 \leq 1$. Perhaps $|PM| = 1 \in \mathbb{Z}$, so the line passing through P and M could be tangent to the hypersphere $x^2 + y^2 + z^2 + t^2 = n$.

We now return to the seemingly strange stuff happening with Theorem 1.11, infinite slopes and the exceptional cases $bm^2 + a = 0$. The problem is that one arrives at the given formula by looking at lines of the form $y = mx+b$ passing through a point (x_0, y_0) on a plane conic, which does not allow for vertical lines (and meets the conic in only one point (x_0, y_0) if $bm^2 + a = 0$). We now give a reformulation of this theorem that treats all lines passing through (x_0, y_0) equally (so there is no issue with an infinite slope being rational). To do this we will use projective geometry and our knowledge of quadratic forms. Instead of searching for rational solutions to $ax^2 + by^2 = c$, we will look for rational solutions to $ax^2 + by^2 = cz^2$ in the projective plane. There is no need to work only over $\mathbb{Q}$.

Let V be a 3-dimensional vector space over a field F (not of characteristic 2) and $Q : V \to F$ be a non-degenerate quadratic form on V. Set

$$C(F) = \{[v] \in \mathbb{P}(V) : Q(v) = 0\}.$$

Theorem A.4. *If $C(F) \neq \emptyset$ then there is a bijection $C(F) \to \mathbb{P}^1(F)$.*

Proof. If $C(F) \neq \emptyset$ then Q has a null vector in V. Write $V = H \perp U$ where H is a hyperbolic plane and $\dim U = 1$. Choose a basis $\{e_1, e_2\}$ of H such that $Q|_H = xy$ in this basis. Extend this basis of H to a basis $\{e_1, e_2, e_3\}$ of V such that

$$Q(xe_1 + ye_2 + ze_3) = xy + \alpha z^2$$

where $\alpha = Q(e_3) \neq 0$. Scale e_1 by $-\alpha$ so

$$Q(x(-\alpha e_1) + ye_2 + ze_3) = -\alpha(xy - z^2).$$

Thus $C(F)$ is the set of all non-trivial solutions in $\mathbb{P}(V)$ to the equation $xy - z^2 = 0$. For this reason, we may consider V to be F^3 and Q to be $xy - z^2$ from now on.

The map $f : \mathbb{P}^1(F) \to C(F) \subset \mathbb{P}^2(F)$ by $[u, v] \mapsto [u^2, v^2, uv]$ is a bijection. To verify this, note that if $[x, y, z] \in C(F)$ then $x \neq 0$ or $y \neq 0$ and define a map $g : C(F) \to \mathbb{P}^1(F)$ by

$$[x, y, z] \mapsto \begin{cases} [x, z], & \text{if } x \neq 0, \\ [z, y], & \text{if } y \neq 0. \end{cases}$$

This is well defined since when $x \neq 0$ and $y \neq 0$, $z \neq 0$ and

$$[x, z] = [1, z/x] = [1, y/z] = [z, y].$$

The maps f and g are mutual inverses so we have a bijection between $C(F)$ and $\mathbb{P}^1(F)$. □

Remark A.5. In Theorem A.4, one can interpret the bijection as coming from intersecting the conic with lines in $\mathbb{P}(V)$ defined over F containing the F-rational point $[1, 0, 0]$ and writing these lines as $c_2y - c_3z = 0$. When this is done, points in $\mathbb{P}^1(F)$ can be viewed as determining particular lines through $[1, 0, 0]$ in $C(F)$.

Appendix B

Index Formulas

In Chapter 3 we used the fact that for a local field K not of characteristic 2, $[K^\times : \mathrm{N}_{L/K}(L^\times)] = 2$ where L is a quadratic extension of K. The goal of this appendix is to prove this assertion (Theorem 3.48) and to use the same methods to obtain an exact formula for $[K^\times : K^{\times^2}]$.

We first treat the case when L/K is a quadratic extension of a local field with odd residue field characteristic since it can be treated in a simpler way.

Theorem B.1. *Let L/K be a quadratic extension of a local field with odd residue field characteristic. Then $[K^\times : \mathrm{N}_{L/K}(L^\times)] = 2$.*

Proof. We have $K^{\times^2} \subset \mathrm{N}_{L/K}(L^\times) \subset K^\times$. We will show $[K^\times : K^{\times^2}] = 4$ and both containments are strict, so the norm-index is 2. From Theorem 3.27 we know that

$$K^\times = \pi_K^{\mathbb{Z}} \times \mu_{q-1} \times (1 + \mathfrak{m}_K)$$

where π_K is a uniformizer, q is the size of the residue field of K and $\mathfrak{m}_K$ is the maximal ideal of $\mathcal{O}_K$ (i.e., $\mathfrak{m}_K = \pi_K \mathcal{O}_K$). So

$$K^{\times^2} = \pi_K^{2\mathbb{Z}} \times \mu_{q-1}^2 \times (1 + \mathfrak{m}_K)$$

since $(1 + \mathfrak{m}_K)^2 = 1 + \mathfrak{m}_K$ by Hensel's lemma. Thus

$$\begin{aligned} K^\times / K^{\times^2} &\cong \mathbb{Z}/2\mathbb{Z} \times \mu_{q-1}/\mu_{q-1}^2 \\ &= \mathbb{Z}/2\mathbb{Z} \times \mu_{q-1}/\mu_{\frac{q-1}{2}} \\ &\cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \end{aligned}$$

so $[K^\times : K^{\times^2}] = 4$. We now show $\mathrm{N}_{L/K}(L^\times) \neq K^\times$ or $K^{\times^2}$.

Let e be the ramification index and f be the residue field degree for L/K, so $2 = ef$. Thus $e = 1$ or $f = 1$. Note that when L/K is a quadratic extension with odd residue field characteristic

$$1 + \mathfrak{m}_K = (1 + \mathfrak{m}_K)^2 \subset \mathrm{N}_{L/K}(1 + \mathfrak{m}_L) \subset 1 + \mathfrak{m}_K$$

so $\mathrm{N}_{L/K}(1 + \mathfrak{m}_L) = 1 + \mathfrak{m}_K$. (This equality is true for other extensions of local fields where $[L : K]$ is not divisible by the residue field characteristic.)

Case 1: (L/K is unramified, so $e = 1$) In this situation π_K is a uniformizer in L so write

$$L^\times \cong \pi_K^{\mathbb{Z}} \times \mu_{q^2-1} \times (1 + \mathfrak{m}_L).$$

Pick a generator η of μ_{q^2-1}. The non-identity automorphism in $\mathrm{Gal}(L/K)$ is the qth power map on μ_{q^2-1}, so

$$\mathrm{N}_{L/K}(\eta) = \eta\bar{\eta} = \eta\eta^q = \eta^{q+1},$$

which has order $q - 1$. Thus $\eta^{q+1} \in K$ is a generator for μ_{q-1}. This means

$$\mathrm{N}_{L/K}(L^\times) \cong \pi_K^{2\mathbb{Z}} \times \mu_{q-1} \times (1 + \mathfrak{m}_K).$$

So π_K is not in $\mathrm{N}_{L/K}(L^\times)$ but is in $K^\times$. Furthermore, there is some non-square Teichmüller representative in $\mathrm{N}_{L/K}(L^\times)$. So $\mathrm{N}_{L/K}(L^\times) \neq K^\times$ or $K^{\times 2}$.

Case 2: (L/K is ramified, so $e = 2$) In this case, a prime in L has prime norm in K. (Let us see this is characteristic of totally ramified extensions of local fields. Write $\pi_K = \pi_L^e u$ for some $u \in \mathcal{O}_L^\times$. The norm of a unit is a unit since for all $x \in L$, $|x| = |\mathrm{N}_{L/K}(x)|^{\frac{1}{n}}$ where $n = [L : K]$. Thus,

$$\pi_K^n = \mathrm{N}_{L/K}(\pi_L^e u) = \mathrm{N}_{L/K}(\pi_L)^e \,\mathrm{N}_{L/K}(u)$$

so $n = e \,\mathrm{ord}_K(\mathrm{N}_{L/K}(\pi_L))$. That is,

$$\mathrm{ord}_K(\mathrm{N}_{L/K}(\pi_L)) = \frac{n}{e} = f,$$

so $\mathrm{N}_{L/K}(\pi_L)$ is a prime in K precisely when L/K is totally ramified.) Furthermore, the residue fields of K and L are isomorphic since $f = 1$. Writing

$$L^\times = \pi_L^{\mathbb{Z}} \times \mu_{q-1} \times (1 + \mathfrak{m}_L),$$

in the quadratic case

$$\mathrm{N}_{L/K}(L^\times) = \mathrm{N}_{L/K}(\pi_L)^{\mathbb{Z}} \times \mu_{q-1}^2 \times (1 + \mathfrak{m}_K).$$

A non-square from $K^\times$ which is in $\mathrm{N}_{L/K}(L^\times)$ is the prime $\mathrm{N}_{L/K}(\pi_L)$. Meanwhile, $\mathrm{N}_{L/K}(L^\times)$ does not contain any non-square Teichmüller representatives from K. Thus $\mathrm{N}_{L/K}(L^\times) \neq K^\times$ or $K^{\times^2}$. □

Remark B.2. In the proof of Theorem B.1 we saw that $[K^\times : K^{\times^2}] = 4$ for all local fields of odd residue field characteristic. We will address the case of 2-adic fields in the last theorem of this appendix.

We now wish to show that Theorem 3.48 is also true for 2-adic fields. To do so we use the p-adic exponential and logarithm, Tate cohomology (for cyclic groups), Herbrand's theorem, and a few results from Galois theory. We will prove Theorem 3.48 for all local fields of characteristic 0 by this method.

We first discuss the p-adic exponential and logarithm. Let K be a local field of characteristic 0 with $\mathcal{O}_K$ as its integer ring and π_K as a uniformizer. There is a descending chain of neighborhoods of 0 in $\mathcal{O}_K$ and 1 in $\mathcal{O}_K^\times$ such that each of the neighborhoods is an open subgroup:

$$\mathcal{O}_K \supset \pi_K\mathcal{O}_K \supset \pi_K^2\mathcal{O}_K \supset \cdots \supset \{0\}$$

and

$$\mathcal{O}_K^\times \supset 1 + \pi_K\mathcal{O}_K \supset 1 + \pi_K^2\mathcal{O}_K \supset \cdots \supset \{1\}.$$

Although $\mathcal{O}_K^\times$ and $\mathcal{O}_K$ are not isomorphic groups ($-1 \in \mathcal{O}_K^\times$ has order 2 whereas there are no non-zero elements of finite order in $\mathcal{O}_K$), we will show that the subgroups $\pi_K^r\mathcal{O}_K$ and $1+\pi_K^r\mathcal{O}_K$ are isomorphic for sufficiently large r using the p-adic exponential and logarithm.

Let K be a finite extension of $\mathbb{Q}_p$ (Theorem 3.24) and normalize the absolute value on K so that $|p| = 1/p$.

Define the formal power series

$$e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}$$

and

$$\log(1+X) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{X^n}{n}.$$

See [3, Section 4.5] for proofs of the following four lemmas.

Lemma B.3. *For $x \in K$, e^x converges if and only if $|x| < p^{-1/(p-1)}$ whereas $\log(1+x)$ converges if and only if $|x| < 1$.*

So we define the p-adic exponential and logarithm on the discs of convergence in K.

Definition B.4. The p-adic exponential on K is

$$\exp : \left\{x \in K : |x| < p^{-1/(p-1)}\right\} \to K \text{ by } x \mapsto e^x.$$

Definition B.5. The p-adic logarithm on K is

$$\log : \{x \in K : |x-1| < 1\} \to K \text{ by } x \mapsto \log(x).$$

Notice that the exponential function on K is defined on an additive group (some $\pi_K^r \mathcal{O}_K$) while the logarithm on K is defined on a multiplicative group $1 + \pi_K \mathcal{O}_K$.

Lemma B.6. *For $x, y \in K$, if $|x|, |y| < p^{-1/(p-1)}$ then*

1) $\exp(x+y) = \exp(x)\exp(y)$,

2) $|\exp(x) - 1| = |x| < p^{-1/(p-1)}$.

Thus $|\exp(x)| = 1$ whenever $|x| < p^{-1/(p-1)}$. So for $|x|, |y| < p^{-1/(p-1)}$,

$$\begin{aligned} |\exp(x) - \exp(y)| &= |\exp(y)||\exp(x-y) - 1| \\ &= |\exp(x-y) - 1| \\ &= |x - y|. \end{aligned}$$

Lemma B.7. *For $x, y \in K$,*

1) *if $|x-1|, |y-1| < 1$ then $\log(xy) = \log(x) + \log(y)$,*

2) *if $|x-1| < p^{-1/(p-1)}$ then $|\log(x)| = |x-1| < p^{-1/(p-1)}$.*

Similarly, for $|x-1|, |y-1| < p^{-1/(p-1)}$,

$$\begin{aligned} |\log(x) - \log(y)| &= |\log(x/y)| \\ &= |x/y - 1| \\ &= |x - y||1/y| \\ &= |x - y|. \end{aligned}$$

Remark B.8. Part 1 of Lemmas B.6 and B.7 show that the p-adic exponential and logarithm functions are homomorphisms while the second part of each of these lemmas shows that they are isometries near the identity (0 or 1). Be aware, however, that the second part of Lemma B.7 does not hold on the whole disc of convergence (i.e., $|x-1|<1$). If x is a p-th root of unity other than 1 then $\log(x^p)=\log(1)$ and $\log(x^p)=p\log(x)$ so $|\log(x)|=0$ but $|x-1|\neq 0$.

Lemma B.9. *Let $x\in K$. If $|x|<p^{-1/(p-1)}$ then $\log(\exp(x))=x$. If $|x-1|<p^{-1/(p-1)}$ then $\exp(\log(x))=x$.*

Theorem B.10. *For sufficiently large r, the groups $\pi_K^r\mathcal{O}_K$ and $1+\pi_K^r\mathcal{O}_K$ are isomorphic.*

Proof. From Lemma B.9 we have that

$$\exp:\{x\in K:|x|<p^{-1/(p-1)}\}\to\{x\in K:|x-1|<p^{-1/(p-1)}\}$$

and

$$\log:\{x\in K:|x-1|<p^{-1/(p-1)}\}\to\{x\in K:|x|<p^{-1/(p-1)}\}$$

are inverse functions. Combining this with Lemmas B.6 and B.7 gives that they are also isometric isomorphisms. Choose $r\in\mathbb{Z}^+$ large enough so that $|\pi_K^r|<p^{-1/(p-1)}$. Then $\exp(\pi_K^r\mathcal{O}_K)=1+\pi_K^r\mathcal{O}_K$. □

We now move on to our discussion of Tate cohomology and Herbrand's theorem.

Definition B.11. Let G be a finite group. A *G-module* is an abelian group A on which G acts by homomorphisms. That is,

1) $1\cdot a=a$ for all $a\in A$,

2) $\sigma(\tau a)=(\sigma\tau)(a)$ for all $\sigma,\tau\in G$ and $a\in A$,

3) $\sigma(a+b)=\sigma a+\sigma b$ for all $\sigma\in G$ and $a\in A$.

Remark B.12. In Definition B.11 we wrote A as a group under addition as a convention only. In practice, A can be a group under multiplication too.

Example B.13. We will keep the following three examples running throughout the discussion:

- (trivial) A is arbitrary and G acts trivially on A: $\sigma a = a$ for all $\sigma \in G$,
- (Galois) K is a field and L/K is a finite Galois extension, $A = L$ or $A = L^\times$ and $G = \mathrm{Gal}(L/K)$,
- (Galois local) K is a local field and L/K is a finite Galois extension, $A = \mathcal{O}_L$ or $A = \mathcal{O}_L^\times$ (where $\mathcal{O}_L$ is the integer ring of L) and $G = \mathrm{Gal}(L/K)$. These are G-modules since $|\sigma(a)| = |a|$ for all $\sigma \in \mathrm{Gal}(L/K)$ and $a \in L$ ($|a| = \sqrt[n]{|\mathrm{N}_{L/K}(a)|}$ for all $a \in L$, where n is the degree of the extension). We could also take $A = L$ or $A = L^\times$ as in the previous case.

Remark B.14. In the Galois or Galois local case of Example B.13, the G-modules $L^\times$ and $\mathcal{O}_L^\times$ are instances where A is written multiplicatively, so differences such as $\sigma a - a$ for $\sigma \in G$ and $a \in A$ become ratios $\sigma a/a$ when this occurs.

Definition B.15. For a G-module A, set

$$A^G = \{a \in A : \sigma(a) = a \text{ for all } \sigma \in G\}.$$

Example B.16. Returning to our previous examples,

- (trivial) $A^G = A$,
- (Galois) $L^G = K$ and $L^{\times G} = K^\times$,
- (Galois local) $\mathcal{O}_L^G = \mathcal{O}_K$ and $\mathcal{O}_L^{\times G} = \mathcal{O}_K^\times$.

Theorem B.17. *Let A be a G-module. When G is cyclic and γ is a generator of G,*

$$A^G = \{a \in A : \gamma(a) = a\}.$$

Proof. If $\gamma(a) = a$ then $\gamma^m(a) = a$ for all $m \in \mathbb{Z}$ so $a \in A^G$. Conversely, if $a \in A^G$ then $\gamma(a) = a$. □

Similar to linear algebra, G-modules have *submodules* (if A is a G-module and A' is a subgroup of A that respects the G-action then A' is a submodule) and *quotient modules* (for example, A/A'). In addition, a *map of G-modules* $f : A_1 \to A_2$ is a homomorphism that satisfies $f(\sigma a_1) = \sigma(f(a_1))$ for $\sigma \in G$ and $a \in A_1$. The kernel and the image of such a map are also G-modules.

Example B.18. Let L/K be a Galois extension of local fields and $G = \mathrm{Gal}(L/K)$. Since $|\sigma(x)| = |x|$ for all $x \in L$ and $\sigma \in G$, the valuation map

$$L^\times \stackrel{\mathrm{ord}}{\to} \mathbb{Z}$$

is a G-module map where $\mathbb{Z}$ has the trivial action:

$$\mathrm{ord}(\sigma(x)) = \mathrm{ord}(x) = \sigma(\mathrm{ord}(x)).$$

Definition B.19. Let A be a G-module. Set $I_G(A) =$ the subgroup (of A) spanned by $\sigma a - a$ for $\sigma \in G$ and $a \in A$.

Theorem B.20. *Let A be a G-module. Then $I_G(A)$ is a submodule.*

Proof. For $\sigma, \tau \in G$ and $a \in A$,

$$\begin{aligned} \tau(\sigma a - a) &= \tau\sigma a - \tau a \\ &= (\tau\sigma a - a) - (\tau a - a) \end{aligned}$$

which is in $I_G(A)$. From this one can see that G acts on $I_G(A)$. □

Theorem B.21. *Let A be a G-module. When G is cyclic and γ is a generator,*

$$I_G(A) = \{\gamma a - a : a \in A\}. \tag{B.1}$$

Proof. The right side of (B.1) is a G-submodule and is contained in $I_G(A)$ so it is enough to show that every $\gamma^m a - a$ in $I_G(A)$ is in $\{\gamma a - a : a \in A\}$. Since $\gamma^{-m} a - a = \gamma^m(\gamma^{-m}(-a)) - \gamma^{-m}(-a)$, we can assume $m \geq 1$:

$$\begin{aligned} \gamma^m a - a &= (\gamma^m a - \gamma^{m-1} a) + (\gamma^{m-1} a - \gamma^{m-2} a) + \cdots + (\gamma a - a) \\ &= \gamma b - b \end{aligned}$$

where $b = \gamma^{m-1} a + \gamma^{m-2} a + \cdots + \gamma a + a$ (which is in A). □

Definition B.22. For a G-module A, set $A_G = A/I_G(A)$.

Remark B.23. Do not confuse A^G with A_G. The first is the largest submodule of A on which G acts trivially while the second is the largest quotient module of A on which G acts trivially ($\sigma a \equiv a \bmod I_G(A)$ so $\sigma(\bar{a}) = \bar{a}$ in A_G). To see that A_G is the largest such quotient module of A, suppose $B \subset A$ is a submodule such that G acts trivially on the quotient A/B. Then $\sigma a \equiv a \bmod B$ for all $\sigma \in G$ and $a \in A$. So $\sigma a - a \in B$ and hence $I_G(A) \subset B$. This means that there is a natural surjective map of G-modules: $f : A/I_G(A) \to A/B$ defined by

$$a \bmod I_G(A) \mapsto a \bmod B.$$

Definition B.24. For a G-module A, define the function $N_G : A \to A^G$ by

$$N_G(a) = \sum_{\sigma \in G} \sigma(a).$$

The values of N_G are in A^G because

$$\tau\left(\sum_{\sigma\in G}\sigma(a)\right) = \sum_{\sigma\in G}\tau(\sigma(a)) = \sum_{\sigma\in G}(\tau\sigma)(a) = \sum_{\sigma\in G}\sigma(a)$$

for all $\tau \in G$. Clearly $N_G : A \to A^G$ is a group homomorphism:

$$N_G(a + a') = N_G(a) + N_G(a').$$

Example B.25. In terms of our running examples,

- (trivial) $N_G(a) = ka$ where $k = \#G$,
- (Galois) when $A = L$, $N_G(x) = \sum_{\sigma\in G}\sigma(x) = \mathrm{Tr}_{L/K}(x)$ and when $A = L^\times$, $N_G(x) = \prod_{\sigma\in G}\sigma(x) = \mathrm{N}_{L/K}(x)$,
- (Galois local) if $A = \mathcal{O}_L$ and $x \in A$ then $N_G(x) = \mathrm{Tr}_{L/K}(x)$ which is in $\mathcal{O}_K$ while if $A = \mathcal{O}_L^\times$ and $x \in A$ then $N_G(a) = \mathrm{N}_{L/K}(x)$ which is in $\mathcal{O}_K^\times$.

The function N_G is also related to A_G because $I_G(A) \subset \ker(N_G)$: for all $\tau \in G$,

$$N_G(\tau a) = \sum_{\sigma\in G}\sigma(\tau a) = \sum_{\sigma\in G}\sigma a = N_G(a).$$

So $N_G(\tau a - a) = 0$. Thus we have an induced function $N_G : A_G \to A^G$.

Definition B.26. Let A be a G-module. The *Tate cohomology* groups of A are $\hat{H}^\circ(A) = A^G/N_G(A)$ and $\hat{H}_\circ(A) = \ker(N_G)/I_G(A)$.

Remark B.27. Note that the G-action on $\hat{H}^\circ(A)$ and $\hat{H}_\circ(A)$ is trivial so these are viewed as abelian groups (just as their name indicates).

Example B.28. In terms of our examples of interest:

- (trivial) $\hat{H}^\circ(A) = A/kA$ where $k = \#G$ and $\hat{H}_\circ(A) = \{a \in A : ka = 0\}$ (note $I_G(A) = \{0\}$ since G acts trivially on A).

- (Galois) For $A = L$, $\hat{H}^\circ(L) = K/\operatorname{Tr}_{L/K}(L) = \{0\}$ since $\operatorname{Tr}_{L/K} : L \to K$ is onto while for $A = L^\times$, $\hat{H}^\circ(L^\times) = K^\times/\operatorname{N}_{L/K}(L^\times)$! If G is cyclic and γ is a generator then

$$\hat{H}_\circ(L) = \{x \in L : \operatorname{Tr}_{L/K}(x) = 0\}/\{\gamma x - x : x \in L\}$$

and

$$\hat{H}_\circ(L^\times) = \{x \in L^\times : \operatorname{N}_{L/K}(x) = 1\}/\{\gamma x/x : x \in L^\times\}.$$

- (Galois local) For $A = \mathcal{O}_L$, $\hat{H}^\circ(\mathcal{O}_L) = \mathcal{O}_K/\operatorname{Tr}_{L/K}(\mathcal{O}_L)$ (note that $\operatorname{Tr}_{L/K}(\mathcal{O}_L)$ is not necessarily all of $\mathcal{O}_K$ so $\hat{H}^\circ(\mathcal{O}_L)$ does not have to be trivial) and for $A = \mathcal{O}_L^\times$, $\hat{H}^\circ(\mathcal{O}_L^\times) = \mathcal{O}_K^\times/\operatorname{N}_{L/K}(\mathcal{O}_L^\times)$. If G is cyclic with a generator γ then

$$\hat{H}_\circ(\mathcal{O}_L) = \{x \in \mathcal{O}_L : \operatorname{Tr}_{L/K}(x) = 0\}/\{\gamma x - x : x \in \mathcal{O}_L\}$$

and

$$\hat{H}_\circ(\mathcal{O}_L^\times) = \{x \in \mathcal{O}_L^\times : \operatorname{N}_{L/K}(x) = 1\}/\{\gamma x/x : x \in \mathcal{O}_L\}.$$

Now is one of the instances where we bring in Galois theory.

Theorem B.29 (Hilbert's Theorem 90)**.** *Let L/K be a cyclic extension, with $G = \operatorname{Gal}(L/K) = \langle\gamma\rangle$. Then, for $x \in L$,*

$$\operatorname{Tr}_{L/K}(x) = 0 \text{ if and only if } x = \gamma y - y \text{ for some } y \in L$$

and

$$\operatorname{N}_{L/K}(x) = 1 \text{ if and only if } x = \frac{\gamma y}{y} \text{ for some } y \in L^\times.$$

Proof. See [6, §6, Chapter VI]. □

So in the Galois case where L is a cyclic extension of K, $\hat{H}_\circ(L)$ and $\hat{H}_\circ(L^\times)$ are trivial by Theorem B.29. Be cautioned, however, that in the Galois local case Theorem B.29 does not imply that $\hat{H}_\circ(\mathcal{O}_L)$ and $\hat{H}_\circ(\mathcal{O}_L^\times)$ are trivial. For instance, if $x \in \mathcal{O}_L$ and $\operatorname{Tr}_{L/K}(x) = 0$ then $x = \gamma y - y$ for some $y \in L$ but why should $y \in \mathcal{O}_L$? We will come back to this later.

We now want to show that the size of each of the cohomology groups $\hat{H}^\circ(L^\times)$, $\hat{H}^\circ(\mathcal{O}_L)$ and $\hat{H}^\circ(\mathcal{O}_L^\times)$ is finite when L/K is an extension of local fields in characteristic 0. Let $n = [L : K]$. Then

$$K^{\times^n} \subset \operatorname{N}_{L/K}(L^\times) \subset K^\times,$$

$$n\mathcal{O}_K \subset \mathrm{Tr}_{L/K}(\mathcal{O}_L) \subset \mathcal{O}_K,$$
$$\mathcal{O}_K^{\times^n} \subset \mathrm{N}_{L/K}(\mathcal{O}_L^\times) \subset \mathcal{O}_K^\times.$$

Furthermore, $[K^\times : K^{\times^n}]$ and $[\mathcal{O}_K^\times : \mathcal{O}_K^{\times^n}]$ are finite by Hensel's lemma while $\mathcal{O}_K/n\mathcal{O}_K$ is clearly finite (K is a local field). Thus $\hat{H}^\circ(L^\times)$, $\hat{H}^\circ(\mathcal{O}_L)$ and $\hat{H}^\circ(\mathcal{O}_L^\times)$ are all finite as we claimed. It is essential that the characteristic of L be 0 because if it had positive characteristic, say p, and $p|n$ then this argument is just nonsense. We will return to this during the proof of Theorem 3.48.

Theorem B.30. *Let A, B be G-modules and $f : A \to B$ be a map of G-modules. Then*

1) $f(A^G) \subset B^G$,

2) $f(N_G(a)) = N_G(f(a))$ *for all* $a \in A$,

3) $f(I_G(A)) \subset I_G(B)$.

Proof. 1) Suppose $a \in A^G$. Then $f(a) = f(\sigma(a)) = \sigma(f(a))$ for all $\sigma \in G$. So $f(a) \in B^G$.

2) By the definition of a map of G-modules we have

$$\begin{aligned} f(N_G(a)) &= f(\sum_{\sigma\in G} \sigma(a)) \\ &= \sum_{\sigma\in G} f(\sigma(a)) \\ &= \sum_{\sigma\in G} \sigma(f(a)) \\ &= N_G(f(a)) \end{aligned}$$

for any $a \in A$.

3) Let $\sigma \in G$ and $a \in A$. Then

$$\begin{aligned} f(\sigma(a) - a) &= f(\sigma(a)) - f(a) \\ &= \sigma(f(a)) - f(a) \end{aligned}$$

so $f(\sigma(a) - a) \in I_G(B)$. This suffices since f is a G-module map. □

Corollary B.31. *Let A, B be G-modules and $f : A \to B$ be a G-module map. Then f induces group homomorphisms:*

$$f^\circ : \hat{H}^\circ(A) \to \hat{H}^\circ(B)$$

and

$$f_\circ : \hat{H}_\circ(A) \to \hat{H}_\circ(B)$$

by $f^\circ(\bar{a}) = \overline{f(a)}$ and $f_\circ(\bar{a}) = \overline{f(a)}$. If $g : B \to C$ is another G-module map then $(gf)^\circ = g^\circ f^\circ$ and $(gf)_\circ = g_\circ f_\circ$.

Proof. By definition $\hat{H}^\circ(A) = A^G/N_G(A)$ and $\hat{H}_\circ(A) = \ker(N_{G,A})/I_G(A)$, where for the purposes of this proof we introduce the notation $N_{G,A}$ for the function $N_G : A \to A$. We first show that f induces a homomorphism from $\hat{H}^\circ(A)$ to $\hat{H}^\circ(B)$. Let a_1, a_2 be in A^G. Suppose $a_1 \equiv a_2 \bmod N_G(A)$. Then

$$f(a_1 - a_2) = f(a_1) - f(a_2)$$

is in $N_G(B)$ by Theorem B.30. So the map $f^\circ : \hat{H}^\circ(A) \to \hat{H}^\circ(B)$ defined by $\bar{a} \mapsto \overline{f(a)}$ is well defined. Furthermore, it is a homomorphism because f is a homomorphism.

We now show that f induces a map from $\hat{H}_\circ(A)$ to $\hat{H}_\circ(B)$. By Theorem B.30, $f(I_G(A)) \subset I_G(B)$ and $f(N_{G,A}(a)) = N_{G,B}(f(a))$ for all $a \in A$. So if a is in $\ker(N_{G,A})$ then $f(a)$ is in $\ker(N_{G,B})$. Hence $f(\ker(N_{G,A})) \subset \ker(N_{G,B})$ and as a result $f_\circ : \hat{H}_\circ(A) \to \hat{H}_\circ(B)$ by $\bar{a} \mapsto \overline{f(a)}$ is well-defined and a homomorphism (again because f is additive).

The last assertion, that $(gf)^\circ = g^\circ f^\circ$ and $(gf)_\circ = g_\circ f_\circ$, is now trivially true. □

Theorem B.32. *If*

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

is an exact sequence of G-modules then there are induced group homomorphisms

$$\hat{H}^\circ(A) \xrightarrow{f^\circ} \hat{H}^\circ(B) \xrightarrow{g^\circ} \hat{H}^\circ(C) \tag{B.2}$$

and

$$\hat{H}_\circ(A) \xrightarrow{f_\circ} \hat{H}_\circ(B) \xrightarrow{g_\circ} \hat{H}_\circ(C) \tag{B.3}$$

such that $g^\circ f^\circ = 0$, $g_\circ f_\circ = 0$ and (B.2) *and* (B.3) *are exact at $\hat{H}^\circ(B)$ and $\hat{H}_\circ(B)$.*

Proof. Use Corollary B.31 to get the induced homomorphisms so $g^\circ f^\circ = 0$ and $g_\circ f_\circ = 0$ because $g \circ f = 0$. This means that $\mathrm{im}(f^\circ) \subset \ker(g^\circ)$ and $\mathrm{im}(f_\circ) \subset \ker(g_\circ)$. We show the reverse implications. To show that $\ker(g^\circ) \subset \mathrm{im}(f^\circ)$, suppose that $g^\circ(\bar{b}) = 0$ for $\bar{b} \in \hat{H}^\circ(B)$. Then $g(b) \in N_G(C)$ so $g(b) = N_G(c)$ for some $c \in C$. But g is onto so $c = g(b')$ for some $b' \in B$. Thus

$$g(b) = N_G(g(b')) = g(N_G(b'))$$

and consequently $b - N_G(b')$ is in $\ker(g) = \mathrm{im}(f)$. This shows that

$$b - N_G(b') = f(a) \tag{B.4}$$

for some $a \in A$. So $f(a) \in B^G$ since $b, N_G(b')$ are in B^G. This implies $\sigma(f(a)) = f(\sigma(a)) = f(a)$ so $\sigma(a) = a$ for all $\sigma \in G$ because f is injective. Hence $a \in A^G$ and we get

$$\bar{b} = f^\circ(\bar{a})$$

when (B.4) is viewed in $\hat{H}^\circ(B)$. Therefore, $\ker(g^\circ) \subset \mathrm{im}(f^\circ)$.

To show that $\ker(g_\circ) \subset \mathrm{im}(f_\circ)$, suppose that $g_\circ(\bar{b}) = 0$ for $\bar{b} \in \hat{H}_\circ(B)$. Then $g(b) \in I_G(C)$ so write $g(b) = \sum_{\sigma\in G} m_\sigma(\sigma(c_\sigma) - c_\sigma)$ for $m_\sigma \in \mathbb{Z}$ and $c_\sigma \in C$. Since g is onto we have that $c_\sigma = g(b_\sigma)$ for some $b_\sigma \in B$. Thus

$$g(b) = \sum_{\sigma\in G} m_\sigma(\sigma(g(b_\sigma)) - g(b_\sigma)) = g(\sum_{\sigma\in G} m_\sigma(\sigma(b_\sigma) - b_\sigma))$$

so $b - \sum_{\sigma\in G} m_\sigma(\sigma(b_\sigma) - b_\sigma))$ is in $\ker(g) = \mathrm{im}(f)$. This implies that

$$b - \sum_{\sigma\in G} m_\sigma(\sigma(b_\sigma) - b_\sigma)) = f(a) \tag{B.5}$$

for some $a \in A$. Moreover, $0 = N_G(f(a)) = f(N_G(a))$ so $N_G(a) = 0$ since f is one-to-one. Viewing (B.5) in terms of $\hat{H}_\circ(B)$ gives

$$\bar{b} = f_\circ(\bar{a}),$$

and therefore $\ker(g_\circ) \subset \mathrm{im}(f_\circ)$. □

Remark B.33. For G-modules A and B, if $f : A \to B$ is onto, this does not mean f° or $f_\circ$ needs to be onto. For instance, if L/K is a Galois extension of local fields of degree k with $G = \mathrm{Gal}(L/K)$ and we view $L^\times$ as a G-module

in the usual way and $\mathbb{Z}$ as a trivial G-module, the map $L^\times \overset{\text{ord}}{\to} \mathbb{Z}$ is onto, but $\hat{H}^\circ(L^\times) \overset{\text{ord}^\circ}{\to} \hat{H}^\circ(\mathbb{Z}) = \mathbb{Z}/k\mathbb{Z}$ is not necessarily onto for $k > 1$ since

$$\#\hat{H}^\circ(L^\times) = [K^\times : \mathrm{N}_{L/K}(L^\times)] < [L : K] = \#(\mathbb{Z}/k\mathbb{Z})$$

when $G = \mathrm{Gal}(L/K)$ is non-abelian. Similarly, neither f° nor $f_\circ$ needs to be injective if f is injective.

Theorem B.34. *Let G be cyclic. When*

$$0 \to A \overset{f}{\to} B \overset{g}{\to} C \to 0$$

is an exact sequence of G-modules, homomorphisms $\delta : \hat{H}_\circ(C) \to \hat{H}^\circ(A)$ and $\delta' : \hat{H}^\circ(C) \to \hat{H}_\circ(A)$ exist such that we have an exact rectangle:

$$\begin{array}{ccccc} \hat{H}^\circ(A) & \xrightarrow{f^\circ} & \hat{H}^\circ(B) & \xrightarrow{g^\circ} & \hat{H}^\circ(C) \\ \uparrow{\scriptstyle \delta} & & & & \downarrow{\scriptstyle \delta'} \\ \hat{H}_\circ(C) & \xleftarrow[g_\circ]{} & \hat{H}_\circ(B) & \xleftarrow[f_\circ]{} & \hat{H}_\circ(A). \end{array}$$

Proof. We first define $\delta : \hat{H}_\circ(C) \to \hat{H}^\circ(A)$. Let $\bar{c} \in \hat{H}_\circ(C)$, so $c \in \ker(N_G)$. Then $c = g(b)$ for some $b \in B$ because g is onto. Thus

$$\begin{aligned} 0 &= N_G(c) \\ &= N_G(g(b)) \\ &= g(N_G(b)) \end{aligned}$$

so $N_G(b) \in \ker(g) = \mathrm{im}(f)$. Since f is one-to-one we have $N_G(b) = f(a)$ for a unique $a \in A$ (the uniqueness is needed for showing δ to be a homomorphism). This means $f(a) \in N_G(B) \subset B^G$. That is,

$$\begin{aligned} f(a) &= \sigma(f(a)) \\ &= f(\sigma(a)) \end{aligned}$$

for all $\sigma \in G$. So $\sigma(a) = a$ (and $a \in A^G$) since f is injective. Hence there is a class $\bar{a}$ of $\hat{H}^\circ(A)$ attached to $\bar{c} \in \hat{H}_\circ(C)$.

We show that $\delta : \hat{H}_\circ(C) \to \hat{H}^\circ(A)$ by $\delta(\bar{c}) = \bar{a}$ where $f(a) = N_G(b)$ with $g(b) = c$ is a well-defined function. Suppose $c = g(b')$ for $b' \in B$. Then $b - b'$

is in the kernel of g which is the image of f: $b' = b + f(a')$ for some $a' \in A$. This means

$$\begin{aligned} N_G(b') &= N_G(b) + N_G(f(a')) \\ &= f(a) + f(N_G(a')) \\ &= f(a + N_G(a')). \end{aligned}$$

But $a + N_G(a') \equiv a \bmod N_G(A)$ so δ is well-defined. Notice that the cyclicity of G was not used.

We now define $\delta' : \hat{H}^\circ(C) \to \hat{H}_\circ(A)$. Let $\bar{c} \in \hat{H}^\circ(C)$. Write $c = g(b)$ for some $b \in B$ since g is onto. Then for any $\sigma \in G$, $\sigma(c) = c$ so

$$\begin{aligned} g(b) &= \sigma(g(b)) \\ &= g(\sigma(b)). \end{aligned}$$

This implies $\sigma(b) - b$ is in $\ker(g) = \mathrm{im}(f)$. Thus $\sigma(b) - b = f(a)$ for a unique $a \in A$ since f is one-to-one. As σ varies, however, so does a. This is where the cyclicity of G plays a role. Fix a generator γ of G. Run through the same argument with γ in place of σ. So $c = g(b)$ and

$$\gamma(b) - b = f(a)$$

for a unique $a \in A$. Furthermore, we have $f(N_G(a)) = N_G(f(a)) = 0$ because $f(a) = \gamma(b) - b$ which is in $I_G(B) \subset \ker(N_G(B))$. Thus $N_G(a) = 0$ since f is injective. So for $\bar{c} \in \hat{H}^\circ(C)$ there is a class $\bar{a} \in \hat{H}_\circ(A)$ connected to it.

We show that $\delta' : \hat{H}^\circ(C) \to \hat{H}_\circ(A)$ by $\bar{c} \mapsto \bar{a}$ where $f(a) = \gamma(b) - b$ with $g(b) = c$ is well-defined. Suppose $c = g(b')$ for $b' \in B$. Then $b' - b$ is in the kernel of g and hence the image of f. Thus $b' = b + f(a')$ for some $a' \in A$. So $\gamma(b') = \gamma(b) + f(\gamma(a'))$ gives

$$\begin{aligned} \gamma(b') - b' &= \gamma(b) - b + f(\gamma(a')) - f(a') \\ &= f(a) + f(\gamma(a')) - f(a') \\ &= f(a + \gamma(a') - a'). \end{aligned}$$

But $a + \gamma(a') - a' \equiv a \bmod I_G(A)$ so δ' is well-defined. Notice how the cyclicity of G played a major role in the definition of δ'. This means δ' depends on a choice of generator. Showing that δ and δ' are homomorphisms and the exactness of the rectangle are left to the reader. $\square$

Definition B.35. For a G-module A such that $\hat{H}^\circ(A)$ and $\hat{H}_\circ(A)$ are finite, its *Herbrand quotient* is $\mathrm{h}(A) = \#\hat{H}^\circ(A)/\#\hat{H}_\circ(A)$.

Example B.36. The Herbrand quotients of some examples (using Example B.28) are:

- (trivial) If $A = \mathbb{Z}$ and $k = \#G$ then

$$\mathrm{h}(\mathbb{Z}) = \frac{\#\mathbb{Z}/k\mathbb{Z}}{\#\{a \in \mathbb{Z} : ka = 0\}} = k.$$

- (Galois) $\hat{H}^\circ(L)$ is trivial and when G is cyclic, $\hat{H}_\circ(L)$ and $\hat{H}_\circ(L^\times)$ are also trivial. Thus when G is cyclic $\mathrm{h}(L) = 1$ and

$$\mathrm{h}(L^\times) = \#(K^\times/\,\mathrm{N}_{L/K}(L^\times)) = [K^\times : \mathrm{N}_{L/K}(L^\times)]$$

  if this index is finite.

- (Galois local) The indices for the multiplicative groups of fields in the previous example are finite in this case when the local fields do not have characteristic dividing $\#G$.

We now treat Herbrand's theorem and a few of its corollaries.

Theorem B.37 (Herbrand)**.** *Let G be cyclic and*

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

be an exact sequence of G-modules. If any two of $\mathrm{h}(A), \mathrm{h}(B)$, *or* $\mathrm{h}(C)$ *are finite then so is the third and*

$$\mathrm{h}(B) = \mathrm{h}(A)\mathrm{h}(C).$$

Proof. By Theorem B.34 we have the exact rectangle

$$\begin{array}{ccccc}
\hat{H}^\circ(A) & \xrightarrow{f^\circ} & \hat{H}^\circ(B) & \xrightarrow{g^\circ} & \hat{H}^\circ(C) \\
\uparrow{\scriptstyle \delta} & & & & \downarrow{\scriptstyle \delta'} \\
\hat{H}_\circ(C) & \xleftarrow[g_\circ]{} & \hat{H}_\circ(B) & \xleftarrow[f_\circ]{} & \hat{H}_\circ(A)
\end{array}$$

where a cohomology group of one module is sandwiched between cohomology groups of the other two modules. For instance, $\hat{H}^\circ(A)$ is in between $\hat{H}_\circ(C)$ and $\hat{H}^\circ(B)$.

In general, if G_1, G_2, G_3 are abelian groups with G_1 and G_3 finite and ψ, φ are homomorphisms such that the sequence

$$G_1 \xrightarrow{\psi} G_2 \xrightarrow{\varphi} G_3$$

is exact at G_2, then G_2 is finite. To see this consider that $G_2/\ker(\varphi)$ embeds into G_3 so $G_2/\operatorname{im}(\psi)$ also embeds into G_3 ($\operatorname{im}(\psi) = \ker(\varphi)$). Thus we have that $[G_2 : \operatorname{im}(\psi)]$ is finite since G_3 is finite. As a result, $\#G_2$ is finite because $\operatorname{im}(\psi)$ is finite too. Applying this general situation to the exact rectangle shows that as long as two of A, B, and C have finite cohomology groups $\hat{H}^\circ$ and $\hat{H}_\circ$, then $\hat{H}^\circ$ and $\hat{H}_\circ$ are finite for the third module also. This means the Herbrand quotient is defined for A, B, and C.

Label the six cohomology groups $\hat{H}^\circ(A), \hat{H}^\circ(B), \ldots, \hat{H}_\circ(C)$ in order of their appearance in the rectangle, starting with say $\hat{H}^\circ(A)$, by $M_1, M_2, \ldots, M_6$ (it does not matter where we start just as long as the order is correct). We already know that each M_i is finite. Let r_i be the size of the kernel out of M_i and let s_i be the size of the image coming into M_i. Since the rectangle is exact, $r_i = s_i$. Furthermore, $\#M_i = r_i s_{i+1}$ because $\#M_i/r_i = s_{i+1}$ by the exactness at M_i of the rectangle (consider $i \in \mathbb{Z}/6\mathbb{Z}$). Thus

$$\begin{aligned} \#M_2\#M_4\#M_6 &= r_2 s_3 r_4 s_5 r_6 s_7 \\ &= s_2 r_3 s_4 r_5 s_6 r_1 \\ &= \#M_1\#M_3\#M_5. \end{aligned}$$

so $\mathrm{h}(A)\mathrm{h}(C) = \mathrm{h}(B)$ since $\mathrm{h}(A) = \#M_1/\#M_4$, $\mathrm{h}(B) = \#M_2/\#M_5$, and $\mathrm{h}(C) = \#M_3/\#M_6$. □

Corollary B.38. *Let G be cyclic. For any exact sequence of G-modules with finite Herbrand quotients*

$$0 \to A_1 \to A_2 \to \cdots \to A_n \to 0,$$

we have $\prod_{i=1}^n \mathrm{h}(A_i)^{(-1)^i} = 1$.

Proof. For $n = 3$ use Herbrand's theorem. Suppose $n > 3$ and that the corollary is true for any exact sequence of $n-1$ G-modules with finite Herbrand quotients. So consider the exact sequence

$$0 \to A_2/A_1 \to A_3 \to \cdots \to A_n \to 0$$

of $n-1$ G-modules. Since

$$0 \to A_1 \to A_2 \to A_2/A_1 \to 0$$

is exact we have $\mathrm{h}(A_2/A_1) = \mathrm{h}(A_2)/\mathrm{h}(A_1)$ and, by the induction hypothesis,

$$\prod_{i=1}^{n} \mathrm{h}(A_i)^{(-1)^i} = \mathrm{h}(A_2/A_1) \prod_{i=3}^{n} \mathrm{h}(A_i)^{(-1)^i} = 1.$$

□

Theorem B.39. *If $A_1, \ldots, A_n$ are finite abelian groups and*

$$0 \to A_1 \to A_2 \to \cdots \to A_n \to 0$$

is an exact sequence then $\prod_{i=1}^{n} (\#A_i)^{(-1)^i} = 1$.

Proof. Induct as in the proof of Corollary B.38. □

Corollary B.40. *Let G be cyclic. For a finite G-module A, $\mathrm{h}(A) = 1$.*

Proof. The Herbrand quotient $\mathrm{h}(A)$ is finite since $\hat{H}^\circ(A)$ and $\hat{H}_\circ(A)$ are finite (the corollary assumes A to be finite). Pick a generator γ for G. We get estimates on the size of $\mathrm{h}(A)$ by considering the following exact sequences of G-modules:

$$0 \to A^G \to A \to A \to A/I_G(A) \to 0$$

where the beginning map is inclusion, the ending map is reduction and the map from A to A is defined by $a \mapsto \gamma a - a$, and

$$0 \to \hat{H}_\circ(A) \to A/I_G(A) \xrightarrow{N_G} A^G \to \hat{H}^\circ(A) \to 0$$

where again the beginning and ending maps are inclusion and reduction. The first exact sequence gives that $\#A^G = \#(A/I_G(A))$ by Theorem B.39. Then the second exact sequence shows that $\#\hat{H}_\circ(A) = \#\hat{H}^\circ(A)$ since $\#A^G = \#(A/I_G(A))$. Thus

$$\mathrm{h}(A) = \frac{\#\hat{H}^\circ(A)}{\#\hat{H}_\circ(A)} = 1.$$

□

Corollary B.41. *Let G be cyclic. For any G-module A and G-submodule B with finite index, $\mathrm{h}(A) = \mathrm{h}(B)$ if either $\mathrm{h}(A)$ or $\mathrm{h}(B)$ is defined.*

Proof. Consider the exact sequence

$$0 \to B \to A \to A/B \to 0.$$

Corollary B.40 shows that $\mathrm{h}(A/B) = 1$ so $\mathrm{h}(A) = \mathrm{h}(B)$ by Herbrand's theorem. □

Lemma B.42. *Let L be a finite extension of $\mathbb{Q}_p$. For cyclic $G = \mathrm{Gal}(L/K)$ when $\mathbb{Q}_p \subset K \subset L$ and L/K is Galois we have $\mathrm{h}(L^\times) = \mathrm{h}(\mathcal{O}_L^\times)[L : K]$ and $\mathrm{h}(\mathcal{O}_L^\times) = \mathrm{h}(\mathcal{O}_L)$.*

Proof. Consider the exact sequence of G-modules

$$0 \to \mathcal{O}_L^\times \to L^\times \stackrel{\mathrm{ord}}{\to} \mathbb{Z} \to 0$$

where the G-action on $\mathbb{Z}$ is trivial ($\mathrm{ord}(\sigma x) = \mathrm{ord}(x) = \sigma\,\mathrm{ord}(x)$ for all $\sigma \in G$ and $x \in L^\times$). Recall from Example B.36 that $\mathrm{h}(L^\times) = [K^\times : \mathrm{N}_{L/K}(L^\times)]$ and that $\mathrm{h}(\mathbb{Z}) = k$ where $k = \#G = [L : K]$. Thus Herbrand's theorem gives us that $\mathrm{h}(\mathcal{O}_L^\times)$ is finite and

$$\begin{aligned} \mathrm{h}(L^\times) &= \mathrm{h}(\mathcal{O}_L^\times)\mathrm{h}(\mathbb{Z}) \\ &= \mathrm{h}(\mathcal{O}_L^\times)[L : K]. \end{aligned}$$

Let U be any G-submodule of $\mathcal{O}_L^\times$ with finite index. Then $\mathrm{h}(\mathcal{O}_L^\times)$ equals $\mathrm{h}(U)$ using Corollary B.41. Since $|\sigma(\pi_L)| = |\pi_L|$, $\pi_L^r\mathcal{O}_L$ is a G-submodule of $\mathcal{O}_L$ and $1 + \pi_L^r\mathcal{O}_L$ is a G-submodule of $\mathcal{O}_L^\times$. From Theorem B.10 we have the isomorphism

$$1 + \pi_L^r\mathcal{O}_L = \exp(\pi_L^r\mathcal{O}_L) \cong \pi_L^r\mathcal{O}_L$$

as groups for $r \gg 0$ since $\mathbb{Q}_p \subset L$. The exponential function is actually an isomorphism of the G-submodules $1 + \pi_L^r\mathcal{O}_L$ and $\pi_L^r\mathcal{O}_L$ for $r \gg 0$ since for small $x \in L$ and $\sigma \in G$, $\exp(\sigma x) = \sigma(\exp(x))$ (to see this, just write out the series of $\exp(x)$). Hence

$$\begin{aligned} \mathrm{h}(\mathcal{O}_L^\times) &= \mathrm{h}(1 + \pi_L^r\mathcal{O}_L) \\ &= \mathrm{h}(\pi_L^r\mathcal{O}_L) \\ &= \mathrm{h}(\mathcal{O}_L). \end{aligned}$$

□

We state one final result from Galois theory before proceeding with our proof of Theorem 3.48 for local fields of characteristic 0.

Definition B.43. Let L/K be a Galois extension. When there is an $x_0 \in L$ such that $\{\sigma(x_0) : \sigma \in G\}$ is a K-basis of L:

$$L \cong \bigoplus_{\sigma \in G} K\sigma(x_0),$$

we call such a basis a *normal* basis.

Theorem B.44. *For any Galois extension L/K, there is a normal basis.*

Proof. See [6, §13, Chapter VI] for a proof when K is infinite. This will be sufficient for our purposes (local fields of characteristic 0). □

To prove Theorem 3.48 for local fields of characteristic 0 we actually prove a stronger result.

Theorem B.45. *Let L/K be a cyclic extension of local fields with characteristic 0. Then $[K^\times : \mathrm{N}_{L/K}(L^\times)] = [L : K]$.*

Proof. Let $G = \mathrm{Gal}(L/K)$. By Lemma B.42,

$$[K^\times : \mathrm{N}_{L/K}(L^\times)] = \mathrm{h}(L^\times) = \mathrm{h}(\mathcal{O}_L^\times)[L : K],$$

so we want to show that $\mathrm{h}(\mathcal{O}_L^\times) = 1$. Lemma B.42 also gives $\mathrm{h}(\mathcal{O}_L^\times) = \mathrm{h}(\mathcal{O}_L)$. In particular, $\mathrm{h}(\mathcal{O}_L^\times) = \mathrm{h}(M)$ where M is any G-submodule of $\mathcal{O}_L$ with finite index by Corollary B.41.

We now make a particular choice of M. From Theorem B.44 we know that a normal basis of L/K exists so let $\{\sigma(x_0) : \sigma \in G\}$ be a normal basis such that x_0 in L is in the disc of convergence of $\exp(X)$. This is possible because we can always scale x_0 by a high power of p to get it in the disc of convergence of $\exp(X)$ (every σ acts trivially on elements of the bottom field). Set $M = \bigoplus_{\sigma \in G} \mathcal{O}_K \sigma(x_0)$ (this really is a G-module). Notice that M is contained in the disc of convergence of the exponential function since $|\sigma(x)| = |x|$ so $M \cong \exp(M) \subset \mathcal{O}_L^\times$.

Lastly, we compute the sizes of $\hat{H}^\circ(M)$ and $\hat{H}_\circ(M)$. For $\#\hat{H}^\circ(M)$, let $\sum_{\sigma \in G} c_\sigma \sigma(x_0)$ be in M^G. Then all the c_σ's must be equal so

$$\begin{aligned} M^G &= \mathcal{O}_K \sum_{\sigma \in G} \sigma(x_0) \\ &= \mathcal{O}_K \,\mathrm{Tr}_{L/K}(x_0) \\ &\subset \mathrm{Tr}_{L/K}(M) \end{aligned}$$

and $\mathrm{Tr}_{L/K}(M) \subset M^G$. Thus $\mathrm{Tr}_{L/K}(M) = M^G$ and hence

$$\#\hat{H}^\circ(M) = \#(M^G/\,\mathrm{Tr}_{L/K}(M)) = 1.$$

Before computing $\#\hat{H}_\circ(M)$, consider the following. For $x \in L$, if we write

$$x = \sum_{\sigma\in G} c_\sigma \sigma(x_0)$$

for $c_\sigma \in K$ then we have

$$\begin{aligned}\mathrm{Tr}_{L/K}(x) &= \sum_{\sigma\in G} c_\sigma \,\mathrm{Tr}_{L/K}(\sigma(x_0)) \\ &= \mathrm{Tr}_{L/K}(x_0)\sum_{\sigma\in G} c_\sigma.\end{aligned}$$

But $\mathrm{Tr}_{L/K}(L) = K$ so $\mathrm{Tr}_{L/K}(x_0) \neq 0$ since otherwise $\mathrm{Tr}_{L/K}(x) = 0$ for all $x \in L$. Thus $\mathrm{Tr}_{L/K}(x) = 0$ if and only if $\sum_{\sigma\in G} c_\sigma = 0$.

We now compute $\#\hat{H}_\circ(M)$. Let $x \in M$ and $\mathrm{Tr}_{L/K}(x) = 0$. As above, write $x = \sum_{\sigma\in G} c_\sigma\sigma(x_0)$. Then $\sum_{\sigma\in G} c_\sigma = 0$ since $\mathrm{Tr}_{L/K}(x) = 0$. So we have

$$\begin{aligned}x &= \sum_{\sigma\in G} c_\sigma\sigma(x_0) \\ &= \sum_{\sigma\in G} c_\sigma(\sigma(x_0) - x_0) \\ &= \sum_{\sigma\in G}(\sigma(c_\sigma x_0) - c_\sigma x_0)\end{aligned}$$

which is in $I_G(M)$. Thus $\ker(\mathrm{Tr}_{L/K} : M \to M) = I_G(M)$ and hence $\#\hat{H}_\circ(M) = 1$. So $\mathrm{h}(M) = 1 = \mathrm{h}(\mathcal{O}_L^\times)$. □

Remark B.46. Theorem B.45 is true even if the characteristic of the local fields is positive, but a different proof is needed to show this because in our proof of Theorem B.45 we used isomorphic finite-index subgroups of $\mathcal{O}_L$ and $\mathcal{O}_L^\times$. If L has positive characteristic, say p, then every non-identity element of L has order p, but no non-identity elements of $L^\times$ have order p, so $\mathcal{O}_L$ and $\mathcal{O}_L^\times$ do not have isomorphic finite-index subgroups.

Recall that $[K^\times : K^{\times^2}] = 4$ for all local fields with odd residue field characteristic as shown in the proof of Theorem B.1. We now find an explicit

formula for $[K^\times : K^{\times m}]$ where $m \in \mathbb{Z}^+$ and K is any local field of characteristic 0. The next two examples will be used in the calculation of this formula. Instead of Galois actions we use trivial actions.

Example B.47. If G acts trivially on a field L then $\hat{H}^\circ(L^\times) = L^\times/L^{\times k}$ and

$$\hat{H}_\circ(L^\times) = \{a \in L^\times : a^k = 1\} = \mu_k(L),$$

the kth roots of unity in L, where $k = \#G$.

Example B.48. If L is a field on which G acts trivially then

$$\mathrm{h}(L^\times) = \frac{[L^\times : L^{\times k}]}{\#\mu_k(L)},$$

where $k = \#G$ when $L^{\times k}$ has finite index in $L^\times$.

Theorem B.49. *Let K be a local field with* $\mathrm{char}(K) = 0$. *Then for $m \in \mathbb{Z}^+$,*

$$[K^\times : K^{\times m}] = m[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}] = \frac{m}{|m|_p^n}\#\mu_m(K),$$

where $\mu_m(K)$ is the group of mth roots of unity of K.

Proof. We first show $[K^\times : K^{\times m}] = m[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}]$. Let π_K be a uniformizer of K. We know $K^\times \cong \pi_K^{\mathbb{Z}} \times \mathcal{O}_K^\times$ so

$$K^{\times m} \cong \pi_K^{m\mathbb{Z}} \times \mathcal{O}_K^{\times m}.$$

Thus $K^\times/K^{\times m} \cong \mathbb{Z}/m\mathbb{Z} \times \mathcal{O}_K^\times/\mathcal{O}_K^{\times m}$, which has size $m[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}]$.

We now compute $[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}]$. Let $G = \mathbb{Z}/m\mathbb{Z}$ act trivially on K and let $n = [K : \mathbb{Q}_p]$. Consider the Herbrand quotient

$$\mathrm{h}(\mathcal{O}_K^\times) = \frac{[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}]}{\#\mu_m(\mathcal{O}_K)} = \frac{[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}]}{\#\mu_m(K)}.$$

In a similar manner as in the proof of Lemma B.42 (but using trivial actions on K and $K^\times$) we can show that

$$\mathrm{h}(\mathcal{O}_K^\times) = \mathrm{h}(\mathcal{O}_K) = [\mathcal{O}_K : m\mathcal{O}_K].$$

So what is $[\mathcal{O}_K : m\mathcal{O}_K]$? Write $m = p^r m'$ where $(p, m') = 1$. Then

$$\mathcal{O}_K/m\mathcal{O}_K = \mathcal{O}_K/p^r\mathcal{O}_K = \mathcal{O}_K/\pi_K^{er}\mathcal{O}_K$$

where e is the ramification index of K over $\mathbb{Q}_p$. Since we have a filtration

$$\mathcal{O}_K \supset \pi_K \mathcal{O}_K \supset \cdots \supset \pi_K^{er-1} \mathcal{O}_K \supset \pi_K^{er} \mathcal{O}_K$$

and $\pi_K^i \mathcal{O}_K / \pi_K^{i+1} \mathcal{O}_K \cong \mathcal{O}_K / \pi_K \mathcal{O}_K$ as groups (divide by π_K^i),

$$\begin{aligned} [\mathcal{O}_K : \pi_K^{er} \mathcal{O}_K] &= \prod_{i=0}^{er-1} [\pi_K^i \mathcal{O}_K : \pi_K^{i+1} \mathcal{O}_K] \\ &= [\mathcal{O}_K : \pi_K \mathcal{O}_K]^{er} \\ &= (p^f)^{er} \\ &= p^{nr}, \end{aligned}$$

which is $1/|m|_p^n$. Thus

$$[\mathcal{O}_K^\times : \mathcal{O}_K^{\times m}] = \#\mu_m(K) \mathrm{h}(\mathcal{O}_K^\times) = \frac{\#\mu_m(K)}{|m|_p^n}.$$

□

Remark B.50. Theorem B.49 shows that $[K^\times : K^{\times 2}] = 2^{2+[K:\mathbb{Q}_2]}$ when K is a 2-adic field.

Bibliography

[1] Z. I. Borevich and I. R. Shafarevich. *Number Theory*. Academic Press, New York, 1966.

[2] J. W. S. Cassels. *Rational Quadratic Forms*. Academic Press, New York, 1978.

[3] F. Q. Gouvêa. *p-adic Numbers: An Introduction*. Springer-Verlag, New York, second edition, 1997.

[4] N. Jacobson. *Basic Algebra II*. W. H. Freeman and Co., New York, second edition, 1989.

[5] N. Koblitz. *p-adic Numbers, p-adic Analysis and Zeta-Functions*. Springer-Verlag, New York, second edition, 1984.

[6] S. Lang. *Algebra*. Springer-Verlag, New York, revised third edition, 2002.

[7] M. Rosen. *Number Theory in Function Fields*. Springer-Verlag, New York, 2002.

[8] J-P. Serre. *A Course in Arithmetic*. Springer-Verlag, New York, 1973.

[9] T. A. Springer. Note on quadratic forms over algebraic number fields. *Indagationes Mathematicae*, 19:39 – 43, 1957.